

93 845

From: Ramirez, Delia
Sent: Wednesday, June 04, 2003 5:55 PM
To: STIC-Biotech/ChemLib
Subject: case 09/777,566

RECEIVED

JUN -5 2003

(STIC)

Hi,

I would like to request the following search (commercial and interference):

1. a standard search of seq id 1 (DNA) in the nucleic acid databases
2. a standard search of seq id 2 (Protein) in the nucleic acid databases
3. a standard search of seq id 2 (Protein) in the protein databases

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
1911 S. Clark Street, Crystal Mall 1, 10D06, Mail room 10D01
Arlington, VA 22202
(703) 306-0288
delia.ramirez@uspto.gov

Edward Han
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/5/03
Date Completed: 6/13/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: I+I Reverse to NA
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 01/02
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpn** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapn** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 06:51:00 ; Search time 64 Seconds
(Without alignments)
6339,584 Million cell updates/sec

Title: US-09-777-566a-1
Perfect score: 1323
Sequence: 1 atgaagcgcattcaatccc.....atccaccatccatcaactaa 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 682724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1322	99.9	1323	US-09-259-214-1	Sequence 1, Appl1
2	1322	99.9	1323	US-09-318-528-1	Sequence 1, Appl1
3	1322	99.9	1323	US-09-291-931-1	Sequence 1, Appl1
4	1210	91.5	1272	US-08-910-798-1	Sequence 1, Appl1
5	34.4	2.6	4403765	US-09-103-840A-2	Sequence 2, Appl1
6	34.4	2.6	4411529	US-09-103-840A-1	Sequence 1, Appl1
7	34	2.6	998	US-07-885-089B-5	Sequence 5, Appl1
8	33	2.5	2511	US-08-680-326-116	Sequence 116, Appl
9	33	2.5	35100	US-08-770-379-17	Sequence 17, Appl
10	33	2.5	35100	US-08-757-669A-17	Sequence 17, Appl
11	33	2.5	35100	US-09-230-371A-17	Sequence 17, Appl
12	31.8	2.4	4403765	US-09-103-840A-2	Sequence 2, Appl1
13	31.8	2.4	4411529	US-09-103-840A-1	Sequence 1, Appl1
14	31.6	2.4	8460	US-08-469-005A-9	Sequence 9, Appl1
15	31.6	2.4	8519	US-09-261-907-1	Sequence 1, Appl1
16	31.2	2.4	1034	US-09-072-596-312	Sequence 312, Appl
17	31	2.3	3147	US-08-441-430-3	Sequence 3, Appl1
18	31	2.3	4488	US-08-441-430-1	Sequence 1, Appl1
19	30.4	2.3	418	US-09-072-596-302	Sequence 302, Appl
20	30.4	2.3	8252	US-08-046-585-15	Sequence 15, Appl
21	30.4	2.3	8252	US-08-393-703-15	Sequence 15, Appl
22	29.8	2.3	8252	PCT-US93-11721-15	Sequence 15, Appl
23	29.8	2.3	9711	US-08-961-527-167	Sequence 167, Appl
24	29.6	2.2	150	US-08-659-473-8	Sequence 8, Appl1
25	29.6	2.2	1005	US-08-875-062-4	Sequence 4, Appl1
26	29.6	2.2	1521	US-08-496-855A-3	Sequence 3, Appl1
27	29.6	2.2	1521	US-07-938-154-9	Sequence 9, Appl1

C 28	29.6	2.2	1521	5	PCT-US91-02311-9	Sequence 9, Appl1
C 29	29.6	2.2	2448	4	US-08-487-596-13	Sequence 13, Appl
C 30	29.6	2.2	2450	2	US-08-466-589-9	Sequence 9, Appl1
C 31	29.6	2.2	2450	2	US-08-700-636-9	Sequence 9, Appl1
C 32	29.6	2.2	2450	3	US-08-467-574-9	Sequence 9, Appl1
C 33	29.6	2.2	2450	4	US-09-217-345-9	Sequence 9, Appl1
C 34	29.6	2.2	6379	1	US-08-499-215-1	Sequence 1, Appl1
C 35	29.2	2.2	508	3	US-08-669-408B-3	Sequence 3, Appl1
C 36	29.2	2.2	1555	3	US-08-669-408B-9	Sequence 9, Appl1
C 37	29	2.2	1548	2	US-08-762-106-5	Sequence 5, Appl1
C 38	29	2.2	1548	4	US-09-320-774-5	Sequence 5, Appl1
C 39	29	2.2	1581	2	US-08-762-106-6	Sequence 6, Appl1
C 40	29	2.2	1581	4	US-09-320-774-6	Sequence 6, Appl1
C 41	29	2.2	3531	1	US-08-629-600-1	Sequence 1, Appl1
C 42	29	2.2	3531	4	US-09-076-761-1	Sequence 1, Appl1
C 43	29	2.2	7680	4	US-09-210-748A-3	Sequence 3, Appl1
C 44	29	2.2	56516	2	US-08-996-306-1	Sequence 1, Appl1
C 45	29	2.2	56516	4	US-09-338-907-1	Sequence 1, Appl1

ALIGNMENTS

91777-566

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 610719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHRYASE
FILE REFERENCE: DIVER1370-1
CURRENT APPLICATION NUMBER: US/09/259, 214A
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: 08/910, 798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1320)
NAME/KEY: misc-feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match 99.9%; Score 1322; DB 3; Length 1323;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAACGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCATTCGA	60
DB	1	ATGAACGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCATTCGA	60
QY	61	TTGCGTCAAGTGAAGCCGAGAGCTGAAGAGTGTGATTCAGTCTGATGCT	120
DB	61	TTGCGTCAAGTGAAGCCGAGAGCTGAAGAGTGTGATTCAGTCTGATGCT	120
QY	121	GTGCGTCTCCAAACCAAGGCGCAGCACTATGAGATGTCACCCGAGCGATGGCCA	180
DB	121	GTGCGTCTCCAAACCAAGGCGCAGCACTATGAGATGTCACCCGAGCGATGGCCA	180
QY	181	ACCTGCGCGGTAAACTGGTGTGCTGACACCGCGGTGAGTAACTGCTATCTC	240
DB	181	ACCTGCGCGGTAAACTGGTGTGCTGACACCGCGGTGAGTAACTGCTATCTC	240
QY	241	GGACATTAACCAAGCCAGCGCTGTGTAAGCCGAGGATTCGTGGCGAAAGGCGTCCG	300
DB	241	GGACATTAACCAAGCCAGCGCTGTGTAAGCCGAGGATTCGTGGCGAAAGGCGTCCG	300

```

301 CAGCTGCTGAGTGCAGGATTAATGCTGATGTCGACGAGCGTACCCGTAACAGCGCA 360
301 CAGCTGCTGAGTGCAGGATTAATGCTGATGTCGACGAGCGTACCCGTAACAGCGCA 360
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
541 CCGCAAAAGCGGTTTCGGGAATGGAAAGGGGCTTATTTCCGCATCAAACTGTGCG 600
541 CCGCAAAAGCGGTTTCGGGAATGGAAAGGGGCTTATTTCCGCATCAAACTGTGCG 600
601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGCAATTCAGCACTC 660
601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGCAATTCAGCACTC 660
661 AAGGTGAGCGCGGACATGTCATTAACCGGCTGCGGTAAGCTTCGCAATGCTGACG 720
661 AAGGTGAGCGCGGACATGTCATTAACCGGCTGCGGTAAGCTTCGCAATGCTGACG 720
721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCCGAGCGGGGTGGGGAAGTCAAC 780
721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCCGAGCGGGGTGGGGAAGTCAAC 780
781 GATTTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACA 840
781 GATTTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACA 840
841 GATTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACA 840
841 GATTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACA 840
841 CGCAGCGCGAGGTTGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
841 CGCAGCGCGAGGTTGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
901 TTGAGCGCGCATCCACCGCAAAACAGAGCGATGATGCTGATGACATTCACCTTCACTG 960
901 TTGAGCGCGCATCCACCGCAAAACAGAGCGATGATGCTGATGACATTCACCTTCACTG 960
961 TTTATTCGCGGACAGATATCTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGG 1020
961 TTTATTCGCGGACAGATATCTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGG 1020
1021 TTTATTCGCGGACAGATATCTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGG 1020
1021 TTTATTCGCGGACAGATATCTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGG 1020
1021 ACGCTCCCGGTCAGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
1021 ACGCTCCCGGTCAGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
1081 CGTGGCGTAAAGCAATACAGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1081 CGTGGCGTAAAGCAATACAGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1081 CGTGGCGTAAAGCAATACAGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1081 CGTGGCGTAAAGCAATACAGCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
1141 CAGATGCGTGAATAAAACGCGCTGTCAATTAATACGCGCGCGGAGAGAGTGAACCTGAC 1200
1141 CAGATGCGTGAATAAAACGCGCTGTCAATTAATACGCGCGCGGAGAGAGTGAACCTGAC 1200
1141 CAGATGCGTGAATAAAACGCGCTGTCAATTAATACGCGCGCGGAGAGAGTGAACCTGAC 1200
1141 CAGATGCGTGAATAAAACGCGCTGTCAATTAATACGCGCGCGGAGAGAGTGAACCTGAC 1200
1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1201 CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGCTGCTGCTGCTGCTGCTGCTGCTG 1260
1261 ATCGTGAATGAAGACGATACCGCGCTGCGAGTTGAGATCTCAACCATACCATACCATAC 1320
1261 ATCGTGAATGAAGACGATACCGCGCTGCGAGTTGAGATCTCAACCATACCATACCATAC 1320
1321 TAA 1323
1321 TAA 1323
1321 TAA 1323

```

```

RESULT 2
US-09-318-528-1
; Sequence 1, Application US/0918528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; EARLIER FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

```

```

Query Match 99.9%; Score 1322; DB 4; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCGCAATCTCA 60
1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCGCAATCTCA 60
61 TTCGCTCAGAGTGAGCCGAGACGTGAAGCTGAAGTGTGATGTCAGTCTCATGT 120
61 TTCGCTCAGAGTGAGCCGAGACGTGAAGCTGAAGTGTGATGTCAGTCTCATGT 120
61 TTCGCTCAGAGTGAGCCGAGACGTGAAGCTGAAGTGTGATGTCAGTCTCATGT 120
61 TTCGCTCAGAGTGAGCCGAGACGTGAAGCTGAAGTGTGATGTCAGTCTCATGT 120
121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
181 ACCTGGCGGTAATAAAGTGGTGGCTGACACCGCGGCGGCGGCGGCGGCGGCGGCGG 240
181 ACCTGGCGGTAATAAAGTGGTGGCTGACACCGCGGCGGCGGCGGCGGCGGCGGCGG 240
181 ACCTGGCGGTAATAAAGTGGTGGCTGACACCGCGGCGGCGGCGGCGGCGGCGGCGG 240
181 ACCTGGCGGTAATAAAGTGGTGGCTGACACCGCGGCGGCGGCGGCGGCGGCGGCGG 240
241 GGACATTAACCAAGCCAGCGCTGTGTAACCGAGGATTTGCGGCAAAAGGCGTCCG 300
241 GGACATTAACCAAGCCAGCGCTGTGTAACCGAGGATTTGCGGCAAAAGGCGTCCG 300
241 GGACATTAACCAAGCCAGCGCTGTGTAACCGAGGATTTGCGGCAAAAGGCGTCCG 300
241 GGACATTAACCAAGCCAGCGCTGTGTAACCGAGGATTTGCGGCAAAAGGCGTCCG 300
301 CAGTGTGTCAGGTGCGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGCGCAA 360
301 CAGTGTGTCAGGTGCGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGCGCAA 360
301 CAGTGTGTCAGGTGCGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGCGCAA 360
301 CAGTGTGTCAGGTGCGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGCGCAA 360
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
361 GCCTTCGCGCGCGGCTGCGACCTGACTGTCATTAACCGTACATACCAGGAGATACG 420
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
421 TCAGATCCCGATCCGTTATTAATCCCTTAATAAACTGCGTGTGGCAACTGGATTAACG 480
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
481 AACGTGACTGACGGGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTACCGGCGAT 540
541 CCGCAAAAGCGGTTTCGGGAATGGAAAGGGGCTTATTTCCGCATCAAACTGTGCG 600
541 CCGCAAAAGCGGTTTCGGGAATGGAAAGGGGCTTATTTCCGCATCAAACTGTGCG 600

```

```

OY 601 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAAAGCAGGACATTACATCGGAATC 660
    |||||||
DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAAAGCAGGACATTACATCGGAATC 660
OY 661 AAGGTGAGGCGCGACATGTCTCATTTAACGGTGGCGGTAAAGCTCGCATCAATGTCTAGC 720
    |||||||
DB 661 AAGGTGAGGCGCGACATGTCTCATTTAACGGTGGCGGTAAAGCTCGCATCAATGTCTAGC 720
OY 721 GAGATATTTCTCTCGACACAGACAGGAATGCGGAGCCGGGGTGGGAAGATCACC 780
    |||||||
DB 721 GAGATATTTCTCTCGACACAGACAGGAATGCGGAGCCGGGGTGGGAAGATCACC 780
OY 781 GATTACACACAGTGGAAACCTTGGTAAGTTGCATTAAGCGCAATTTATTTGCTACAA 840
    |||||||
DB 781 GATTACACACAGTGGAAACCTTGGTAAGTTGCATTAAGCGCAATTTATTTGCTACAA 840
OY 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCCGTATTTGATTTGATCATGCGAGCG 900
    |||||||
DB 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCCGTATTTGATTTGATCATGCGAGCG 900
OY 901 TTGACAGCCCATCCACCGCAAAACAGGCGTATGGTGTACATTTACCCATTCAGTACG 960
    |||||||
DB 901 TTGACAGCCCATCCACCGCAAAACAGGCGTATGGTGTACATTTACCCATTCAGTACG 960
OY 961 TTTATTTGCCGACACGATCTAATCTGGCAAAATCTGGGCGGACCTGGAGCTCAACTGG 1020
    |||||||
DB 961 TTTATTTGCCGACACGATCTAATCTGGCAAAATCTGGGCGGACCTGGAGCTCAACTGG 1020
OY 1021 ACGCTTCCGGTTCAGCGGATTAACACGCGCCAGGTGTGTAACCTGTTTGAACCGCTGG 1080
    |||||||
DB 1021 ACGCTTCCGGTTCAGCGGATTAACACGCGCCAGGTGTGTAACCTGTTTGAACCGCTGG 1080
OY 1081 CGTCGGCTTAAGGATTAACGCGCAGTGTGATTCAGGTTTGGCTGTCTCCAGACTTACAG 1140
    |||||||
DB 1081 CGTCGGCTTAAGGATTAACGCGCAGTGTGATTCAGGTTTGGCTGTCTCCAGACTTACAG 1140
OY 1141 CAGATGCGGTATTAACGCGCGTGTATTAATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
    |||||||
DB 1141 CAGATGCGGTATTAACGCGCGTGTATTAATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
OY 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGGCGATGTCTTGGCAGGTTTACGCAA 1260
    |||||||
DB 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGGCGATGTCTTGGCAGGTTTACGCAA 1260
OY 1261 ATCGTGAATGAAGACAGCATACCGGCGTGTGATTTGATTCATCAACCATCAACCATCAC 1320
    |||||||
DB 1261 ATCGTGAATGAAGACAGCATACCGGCGTGTGATTTGATTCATCAACCATCAACCATCAC 1320
OY 1321 TAA 1323
    |||
DB 1321 TAA 1323

```

```

RESULT 3
US-09-291-931-1
; Sequence 1, Application us/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291, 931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910, 798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259, 214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA

```

```

; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

```

```

Query Match      99.9%; Score 1322; DB 4; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ATGAAGCATCTTATATCCATTTTATCTCTGATTCCTGTTAACCCCGCAATGTGCA 60
    |||||||
DB 1 ATGAAGCATCTTATATCCATTTTATCTCTGATTCCTGTTAACCCCGCAATGTGCA 60
OY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGTGATTTGTCAGTGTCTATGGT 120
    |||||||
DB 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGTGATTTGTCAGTGTCTATGGT 120
OY 121 GTGCGTCTCAACCAAGGCGACGCACTGATGTCAGAGATGTCAACCCAGAGCATGGCCA 180
    |||||||
DB 121 GTGCGTCTCAACCAAGGCGACGCACTGATGTCAGAGATGTCAACCCAGAGCATGGCCA 180
OY 181 ACCTGCGCGTAAATCTGGGTTGGTGTGACACCGCGNGTGTGATGCTAATTCGCTATCTC 240
    |||||||
DB 181 ACCTGCGCGTAAATCTGGGTTGGTGTGACACCGCGNGTGTGATGCTAATTCGCTATCTC 240
OY 241 GGACATTACCAACGCGCAGCGTGTGTAGCCGACGATGCTGTGCGCAAAAAAGGCGTCCCG 300
    |||||||
DB 241 GGACATTACCAACGCGCAGCGTGTGTAGCCGACGATGCTGTGCGCAAAAAAGGCGTCCCG 300
OY 301 CAGTCTGTGACAGTGTGCGGATTTATGCTGATGTGTGACAGAGCTTACCGGTAAACAGCGCA 360
    |||||||
DB 301 CAGTCTGTGACAGTGTGCGGATTTATGCTGATGTGTGACAGAGCTTACCGGTAAACAGCGCA 360
OY 361 GCGTTGCGCGCGGCGTGGCACCTGTGACCTGTGCAATTAACGTAATCCAGAGCAGATAGC 420
    |||||||
DB 361 GCGTTGCGCGCGGCGTGGCACCTGTGACCTGTGCAATTAACGTAATCCAGAGCAGATAGC 420
OY 421 TCCAGTCCGCGATCCGTTATTTAATCTCTAATAAAGTGGCGTTGCCACTGTGTAACGCG 480
    |||||||
DB 421 TCCAGTCCGCGATCCGTTATTTAATCTCTAATAAAGTGGCGTTGCCACTGTGTAACGCG 480
OY 481 AACGTACTGACGCGATCTCTACAGCAGGAGGAGGTCATTTGCTACTTTTACCGGGCAT 540
    |||||||
DB 481 AACGTACTGACGCGATCTCTACAGCAGGAGGAGGTCATTTGCTACTTTTACCGGGCAT 540
OY 541 CGGCAAAAGGCGGTTTGGCGAATGGAAGGCGTCTTAATTTTCCGCAATCAAACTGTGCG 600
    |||||||
DB 541 CGGCAAAAGGCGGTTTGGCGAATGGAAGGCGTCTTAATTTTCCGCAATCAAACTGTGCG 600
OY 601 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAAAGCAGGACATTACATCGGAATC 660
    |||||||
DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTTTCATTAAAGCAGGACATTACATCGGAATC 660
OY 661 AAGGTGAGGCGCGACATGTCTCATTTAACGGTGGCGGTAAAGCTCGCATCAATGTCTAGC 720
    |||||||
DB 661 AAGGTGAGGCGCGACATGTCTCATTTAACGGTGGCGGTAAAGCTCGCATCAATGTCTAGC 720
OY 721 GAGATATTTCTCTCGACACAGACAGGAATGCGGAGCCGGGGTGGGAAGATCACC 780
    |||||||
DB 721 GAGATATTTCTCTCGACACAGACAGGAATGCGGAGCCGGGGTGGGAAGATCACC 780
OY 781 GATTACACACAGTGGAAACCTTGGTAAGTTGCATTAAGCGCAATTTATTTGCTACAA 840
    |||||||
DB 781 GATTACACACAGTGGAAACCTTGGTAAGTTGCATTAAGCGCAATTTATTTGCTACAA 840
OY 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCCGTATTTGATTTGATCATGCGAGCG 900
    |||||||
DB 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCCGTATTTGATTTGATCATGCGAGCG 900

```

```

QY 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 960
D 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 960
QY 961 TTTATGCGGACACGATTAATCTGGCAAAATCTGGCGGCGGACAGTGGACCTCAACTG 1020
D 961 TTTATGCGGACACGATTAATCTGGCAAAATCTGGCGGCGGACAGTGGACCTCAACTG 1020
QY 1021 ACAGTTCGCGTACAGCGGATACACCGCCGAGGTGGTGAACGTGGTTGAACGCTG 1080
D 1021 ACAGTTCGCGTACAGCGGATACACCGCCGAGGTGGTGAACGTGGTTGAACGCTG 1080
QY 1081 CGTGGCTAAGCCTAACAAGCAGTTCAGTTCGCTGCTGCTTCACAGCTTTACAG 1140
D 1081 CGTGGCTAAGCCTAACAAGCAGTTCAGTTCGCTGCTGCTTCACAGCTTTACAG 1140
QY 1141 CAGATGCGTGAATAAAGCGCGTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
D 1141 CAGATGCGTGAATAAAGCGCGTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
QY 1201 CTGGCAGATGTGAAGACGCAAAATCGCAGGCGATGCTGCTGCTGCGAGGTTTACGCA 1260
D 1201 CTGGCAGATGTGAAGACGCAAAATCGCAGGCGATGCTGCTGCTGCGAGGTTTACGCA 1260
QY 1261 ATCGTGAATGAAGCAGCATACCGCGGTGCACTTTAGATCTCATACCATCACCATCAC 1320
D 1261 ATCGTGAATGAAGCAGCATACCGCGGTGCACTTTAGATCTCATACCATCACCATCAC 1320
QY 1321 TAA 1323
D 1321 TAA 1323

```

```

RESULT 4
: Sequence 1, Application US/08910798
: Patent No. 5876997
: GENERAL INFORMATION:
: APPLICANT: KRETEZ
: TITLE OF INVENTION: NOVEL PHYTASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/910,798
: FILING DATE: August 13, 1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HAILE, PH.D., LISA A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09010/029001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1272 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: PHYTASE

```

```

: FEATURE:
: NAME/KEY:
: LOCATION:
: us-08-910-798-1
Query Match 91.5%, Score 1210, DB 2, Length 1272;
Best Local Similarity 96.1%, Pred. No. 0;
Matches 1272; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 ATGAAGGATCTTAATCCATTTTATCTTTGATTCGATTACCCGCAATCTGCA 60
D 1 ATGAAGGATCTTAATCCATTTTATCTTTGATTCGATTACCCGCAATCTGCA 60
QY 61 TTCGCTAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGCTCATGT 120
D 61 TTCGCTAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGCTCATGT 120
QY 61 TTCGCTAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGCTCATGT 120
D 61 TTCGCTAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATTTGATGCTCATGT 120
QY 121 GTGGGTGCTCCACCAAGCCAGCCACTGATGACAGATGTCACCCAGACGCTAGCA 180
D 121 GTGGGTGCTCCACCAAGCCAGCCACTGATGACAGATGTCACCCAGACGCTAGCA 180
QY 121 GTGGGTGCTCCACCAAGCCAGCCACTGATGACAGATGTCACCCAGACGCTAGCA 180
D 121 GTGGGTGCTCCACCAAGCCAGCCACTGATGACAGATGTCACCCAGACGCTAGCA 180
QY 181 ACCTGGCGGTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATGCGCTATCTC 240
D 181 ACCTGGCGGTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATGCGCTATCTC 240
QY 181 ACCTGGCGGTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATGCGCTATCTC 240
D 181 ACCTGGCGGTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATGCGCTATCTC 240
QY 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGAGGATTTGCGGAAAAAGGGCTGCCG 300
D 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGAGGATTTGCGGAAAAAGGGCTGCCG 300
QY 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGAGGATTTGCGGAAAAAGGGCTGCCG 300
D 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGAGGATTTGCGGAAAAAGGGCTGCCG 300
QY 301 CAGTCTGTCAAGTGTGGATTTATGCTGATGTCACAGAGCTTACCCTTAACAGCGCA 360
D 301 CAGTCTGTCAAGTGTGGATTTATGCTGATGTCACAGAGCTTACCCTTAACAGCGCA 360
QY 301 CAGTCTGTCAAGTGTGGATTTATGCTGATGTCACAGAGCTTACCCTTAACAGCGCA 360
D 301 CAGTCTGTCAAGTGTGGATTTATGCTGATGTCACAGAGCTTACCCTTAACAGCGCA 360
QY 361 GCCTTGGCGGCGGCTGGCACCTGACTGTGCATTAACCGTACACCGAGCAGATACG 420
D 361 GCCTTGGCGGCGGCTGGCACCTGACTGTGCATTAACCGTACACCGAGCAGATACG 420
QY 357 -----CAGAGCAATACG 369
D 357 -----CAGAGCAATACG 369
QY 421 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGG 480
D 421 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGG 480
QY 370 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGG 429
D 370 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGG 429
QY 481 AACGTGACTAGCAGGATCTCTAGCAGAGGCGAGAGGCTCAATGCTGACTTACGGGCAT 540
D 481 AACGTGACTAGCAGGATCTCTAGCAGAGGCGAGAGGCTCAATGCTGACTTACGGGCAT 540
QY 430 AACGTGACTAGCAGGATCTCTAGCAGAGGCGAGAGGCTCAATGCTGACTTACGGGCAT 489
D 430 AACGTGACTAGCAGGATCTCTAGCAGAGGCGAGAGGCTCAATGCTGACTTACGGGCAT 489
QY 541 CGGCAAAAGGGGTTTCGGAACCTGGAAGGAGGCTTATTTTCGCAATCAACTTGTGC 600
D 541 CGGCAAAAGGGGTTTCGGAACCTGGAAGGAGGCTTATTTTCGCAATCAACTTGTGC 600
QY 490 CGGCAAAAGGGGTTTCGGAACCTGGAAGGAGGCTTATTTTCGCAATCAACTTGTGC 549
D 490 CGGCAAAAGGGGTTTCGGAACCTGGAAGGAGGCTTATTTTCGCAATCAACTTGTGC 549
QY 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 660
D 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 660
QY 550 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 609
D 550 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATCGGAATC 609
QY 661 AAGGTGAGCGCGGCAATGCTCTAATTAACGCGGTGCTTAAGCCCTGATCAATGCTAGC 720
D 661 AAGGTGAGCGCGGCAATGCTCTAATTAACGCGGTGCTTAAGCCCTGATCAATGCTAGC 720
QY 610 AAGGTGAGCGCGGCAATGCTCTAATTAACGCGGTGCTTAAGCCCTGATCAATGCTAGC 669
D 610 AAGGTGAGCGCGGCAATGCTCTAATTAACGCGGTGCTTAAGCCCTGATCAATGCTAGC 669
QY 721 GAGATATTTTCTCGCAACAGCAGAGGAAATGCGGAGCGGAGGCTGGGGAAGATACG 780
D 721 GAGATATTTTCTCGCAACAGCAGAGGAAATGCGGAGCGGAGGCTGGGGAAGATACG 780
QY 670 GAGATATTTTCTCGCAACAGCAGAGGAAATGCGGAGCGGAGGCTGGGGAAGATACG 729
D 670 GAGATATTTTCTCGCAACAGCAGAGGAAATGCGGAGCGGAGGCTGGGGAAGATACG 729
QY 781 GATTACACAGAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTGCTACAA 840
D 781 GATTACACAGAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTGCTACAA 840
QY 730 GATTACACAGAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTGCTACAA 789
D 730 GATTACACAGAGTGAACACCTTGTGATTTGATTAACGCGCAATTTTATTGCTACAA 789
QY 841 CGCAGCGCAGAGTGGCGCGCAGCGGCGCAACCCGTTATTGATGATGAGCGAGG 900
D 841 CGCAGCGCAGAGTGGCGCGCAGCGGCGCAACCCGTTATTGATGATGAGCGAGG 900
QY 790 CGCAGCGCAGAGTGGCGCGCAGCGGCGCAACCCGTTATTGATGATGAGCGAGG 849
D 790 CGCAGCGCAGAGTGGCGCGCAGCGGCGCAACCCGTTATTGATGATGAGCGAGG 849
QY 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 960
D 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 960
QY 850 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 909
D 850 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGACATTAACCACTTCAGTACTG 909

```

```

Db      3834515  CCGGGCGTGGCGGCTCTCCGGGGTCCCGACGGGTGGTACCGCGCGTGGT 3834567

RESULT 6
US-09-103-840A-1
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-2007.00
: CURRENT APPLICATION NUMBER: US/09/103.840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: patentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      2.6%; Score 34.4; DB 4; Length 4411529;

```

QY 148 CTGATGCGAGTATGTCACCCCGCAGGCAATGGCCAACTGGCCGGTAAACATGGGTTGGCTG 207
 Db 3848436 CTGGCGGGGGGATCCACCGGTCTGGCCGCGCTCACGAGCAGCCGAAATCATCTGGAGCTC 3848495
 QY 208 ACACCGGGNGGTGGTGTGACATTAATGCGCTATCTCGACATTACCAACGCCAGCGTGGTA 267
 Db 3848496 GCTGCACAGGTAGATGCCCTCTGTGAGCAGCGGTGGCATATACCAACGCCGCGGCATGTC 3848555
 QY 268 GCCGACGAGATTGC-TGGCGAAAAAGGGCTGGCCCACTGGTGAGGTGCGGATTATTGC 326
 Db 3848556 AGCGAAGGCGTGCAGATTCCACACCGCAGTACTGCTTCCGCGCGGACATCACGCG 3848615
 QY 327 TGAATGCGACGACGAGCGTACCCGTAATAACAGCGGAAGCCTTGCGCCGCCGCGCTGG 379
 Db 3848616 CCGGCGGCTGCGCGCTCTCCGCGGTCCCGAGAGGGTGGTCAACCGCCGTGCTGG 3848668
 RESULT 7
 US-07-885-089B-5/c
 : Sequence 5, Application US/07885089B
 : Patent No. 5830995
 : GENERAL INFORMATION:
 : APPLICANT: Shoyab, Mohammed
 : APPLICANT: McDonald, Vicki L.

1 APPLICANT: Plowman, Gregory D.
2 TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
3 TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
4 NUMBER OF SEQUENCES: 37
5 CORRESPONDENCE ADDRESS:
6 ADDRESSEE: Pennie & Edmonds LLP
7 STREET: 1155 Avenue of the Americas
8 CITY: New York
9 STATE: New York
10 COUNTRY: USA
11 ZIP: 10036
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: PatentIn Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/07/885,089B

Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGCCAATCTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGGAATGTAACTTCCTGGCAGCAGCCCTCAAGAGGCGAGAAATA 12975

Qy 519 AATTCCTGACTTTACGGGCGATCGCAACGGCGTTTGGCCGAACGTGAGAGGCTCTTA 578
Db 12974 CCTGTATGTGGCCGTCCTGACCTACCTAGGGGAGTCTTGGCCAGCTTGGCTA 12915

Qy 579 TTTCCGCAATCAACTTGTGCTTAAACGTGAGAAACAGAGCAAGC 627
Db 12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAAACGAGAAAGATCATAAC 12866

RESULT 10
US-08-757-669A-17/c
Sequence 17, Application US/08757669A
Patent No. 6183751
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,669A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-669A-17

Query Match 2.5%; Score 33; DB 4; Length 35100;
Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGCCAATCTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGGAATGTAACTTCCTGGCAGCAGCCCTCAAGAGGCGAGAAATA 12975

Qy 519 AATTCCTGACTTTACGGGCGATCGCAACGGCGTTTGGCCGAACGTGAGAGGCTCTTA 578
Db 12974 CCTGTATGTGGCCGTCCTGACCTACCTAGGGGAGTCTTGGCCAGCTTGGCTA 12915

Qy 579 TTTCCGCAATCAACTTGTGCTTAAACGTGAGAAACAGAGCAAGC 627
Db 12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAAACGAGAAAGATCATAAC 12866

RESULT 11
US-09-230-371A-17/c
Sequence 17, Application US/09230371A
Patent No. 6348586
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 45185-G-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,371A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: PCT/US97/13346
PRIOR FILING DATE: 1997-07-22
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 17
LENGTH: 35100
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-17

Query Match 2.5%; Score 33; DB 4; Length 35100;
Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGCCAATCTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGGAATGTAACTTCCTGGCAGCAGCCCTCAAGAGGCGAGAAATA 12975

Qy 519 AATTCCTGACTTTACGGGCGATCGCAACGGCGTTTGGCCGAACGTGAGAGGCTCTTA 578
Db 12974 CCTGTATGTGGCCGTCCTGACCTACCTAGGGGAGTCTTGGCCAGCTTGGCTA 12915

Qy 579 TTTCCGCAATCAACTTGTGCTTAAACGTGAGAAACAGAGCAAGC 627
Db 12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAAACGAGAAAGATCATAAC 12866

RESULT 12
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-2007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

	Query Match	2.4%	Score 31.8	DB 4	Length 4403765
	Best Local Similarity	47.4%	Pred. No. 67		
	Matches	93	Conservative	0	Mismatches 103
					Indels 0
					Gaps 0
OY	196	CTGGGTTGGCTGCACACCGCGMGTTGGTAGCTAATCGCCTTATCTGGACATTACCAACGC	255		
Db	926394	CTGGCGGTTGTGCACACGGGTGTGGTCTGTGCCCGCGCGTGTGCCCGCTCCAAACCAACAA	926335		
OY	256	CAGCGTCTGTAGACCGACGACGATTTCTGGGGAAAAAGAGGTGGCCGACGTCTGGTAGAGATC	315		
Db	926334	CAGCCCAACCGATTGCCGCCGACGCGCGGTGCGCCCGGGGCGATGTGCCATTTCGCCACTGCG	926275		
OY	316	GGCATTTATGCTGATGTGCAGCAGCGTACCCGCTAAACAGCGCGAAGCCTTGGCGCGCGG	375		
Db	926274	GCCGTTTGGCGCGCGTGGCCGATTCAGCCCGGCCCGCGCGCGCCCGGACACCGCGT	926215		
OY	376	CTGGACCTGACTGTG	391		
Db	926214	CCCGGCCCCCGCGGTG	926119		

```

RESULT 13
US-09-103-840A-1/c
: Sequence 1, Application US/09103840A
: Patent No. 6294328
: GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: TITLE OF INVENTION: TUBERCULOSIS
: FILE REFERENCE: 24366-20007.00
: CURRENT APPLICATION NUMBER: US/09/103,840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 441529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

RESULT 14
 US-08-469-005A-9/c
 : Sequence 9, Application US/08469005A
 : Patent No. 5665874
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: KUHAJDA, FRANCIS P.
 : APPLICANT: PASTERNAK, GARY A.
 :
 : TITLE OF INVENTION: CANCER RELATED ANTIGEN

NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BAKER & BOTS, L.L.P.
 STREET: 1299 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20004-2400
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 1.5
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,005A
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/188,426
 FILING DATE: 24-JAN-1994
 APPLICATION NUMBER: 08/096,908
 FILING DATE: 26-JUL-1993
 APPLICATION NUMBER: 07/917,716
 FILING DATE: 24-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Posorske, Laurence H
 REGISTRATION NUMBER: 34,698
 REFERENCE/DOCKET NUMBER: 062482-0113
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-639-7700
 TELEFAX: 202-639-7890
 TELEX:
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8460 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE:
 ORIGINAL SOURCE:
 FEATURE:
 NAME/KEY: Coding Sequence
 LOCATION: 124...7650
 OTHER INFORMATION:
 US-08-469-005A-9

Query Match	2.4%;	Score 31.6;	DB 1;	Length 8460;
Best Local Similarity	47.9%;	Pred. No. 9.1;		
Matches	91;	Conservative	0;	Mismatches 99; Indels 0; Gaps 0;
OY	712	ATGCTGACGGAGATATTTCTCTCTCCACAAACACACAGGAGATATCCGGACCCGGGTGGGGA	771	
Db	7604	ATGCTGATGAGGAGACTCCAGGCCCTGCGCTCCACAGCCGGTGGGTGTACCTCGATG	7545	
OY	772	AGGATTCACCCGATTTACACACAGTGGAAACACTTTCGTAATTGGCATTAACGCGCAATTTTAT	831	
Db	7544	ACGTGGACGGATTACTTTCCCTCCGATTAACCTGGGAGAGAGTTGTATATCGCGGCCACAGTCC	7485	
OY	832	TTTGCTACAAACGACGCGCAGAGGTTGCCCGGACGCCGCCACCCCGTTAATTGGATTGATC	891	
Db	7484	TGCGCGTAGGGCGCCACCCGCTTTGGCGCGCAGTAGATCAACGTTGCCATGTACTTGGCC	7425	
OY	892	ATGGCAGCGT	901	
Db	7424	TTTGGGTGTG	7415	

RESULT 15
US-09-261-907-1/c
; Sequence 1, Application US/09261907A

THIS PAGE BLANK (USPTO)

RESULT	1
AAC88885	
ID	AAC88885 standard; DNA; 1323 BP.
AC	
XX	AAC88885;
XX	
DT	07-MAR-2001 (first entry)
XX	
DE	Escherichia coli B phytase enzyme nucleotide sequence.
XX	
KW	Escherichia coli B; phytase enzyme; anabolic; phytate digestion; nutrition; ds.
XX	
OS	Escherichia coli.
XX	
PN	WO200071728-A1.
XX	
PD	30-NOV-2000.
XX	
PF	25-MAY-2000; 2000WO-US14846.
XX	
PR	25-MAY-1999; 99US-0318528.
XX	
PA	(DIVE-) DIVERSA CORP.
XX	
P1	Short JM, Kretz KA;
XX	
DR	WPI; 2001-112081/12.
DR	P-PSDB; AAB37892.
XX	
PT	Improving the nutritional value of phytate-containing foodstuffs, using phytase enzymes which catalyze the liberation of inorganic phosphate

from the phytates -
 Claim 2; Fig 1; 147pp; English.

The present sequence encodes a phytase enzyme from *Escherichia coli* B. The enzyme catalyses the liberation of inorganic phosphate from the phytate in phytate-containing foodstuffs and can thus be used to improve the nutritional value of phytate rich ingredients.

Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 1 other:

Query Match 99.9%; Score 1322; DB 22; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCAATCTGA
DB 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCAATCTGA
OY 61 TTCGCTCAGAGTGAACCCGAGCTGAAGCTGAAGTGTGTGATTTGTCAGTGCATGCT
DB 61 TTCGCTCAGAGTGAACCCGAGCTGAAGCTGAAGTGTGTGATTTGTCAGTGCATGCT
OY 121 GTGCGTGTCCCAACCAAGCCGCACTGATGCAAGATGTACCCCGAGCCATGGCCA
DB 121 GTGCGTGTCCCAACCAAGCCGCACTGATGCAAGATGTACCCCGAGCCATGGCCA
OY 181 ACCTGGCCGGTAAACTGGTGTGCTGACACCCGCGNGGTGTGAGCTAATCGCTATCTC
DB 181 ACCTGGCCGGTAAACTGGTGTGCTGACACCCGCGNGGTGTGAGCTAATCGCTATCTC
OY 241 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGCTGCCG
DB 241 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGCTGCCG
OY 301 CAGTGTGCTCAGAGTGTGCGATTTATGCTGATGTCGAGCGGACCGGTAACAGGCGAA
DB 301 CAGTGTGCTCAGAGTGTGCGATTTATGCTGATGTCGAGCGGACCGGTAACAGGCGAA
OY 361 GCGTTGCGCGCGGGCTGGACCTGACTGTGCAATTAACCGTATACCGAGCAGATAGC
DB 361 GCGTTGCGCGCGGGCTGGACCTGACTGTGCAATTAACCGTATACCGAGCAGATAGC
OY 421 TCAGTCCCGATCCGTTATTTATCTCTAAAACTGGCGTTGGCAACTGTGATACCGG
DB 421 TCAGTCCCGATCCGTTATTTATCTCTAAAACTGGCGTTGGCAACTGTGATACCGG
OY 481 AACGTGACTGACGCGATCTCAGCAGGAGGAGGATTAATGCGACTTTACCGGGCAT
DB 481 AACGTGACTGACGCGATCTCAGCAGGAGGAGGATTAATGCGACTTTACCGGGCAT
OY 541 CGGCAAAAGCGGGTTTGGGAACTGGAACGGGTGCTTAATTTTCCGAATCAAACTTGTGC
DB 541 CGGCAAAAGCGGGTTTGGGAACTGGAACGGGTGCTTAATTTTCCGAATCAAACTTGTGC
OY 601 CTTAAACGTGAGAAACAGAGAGAAAGCTGTTTATTAACGCGAGCTTACCATCGAAGTC
DB 601 CTTAAACGTGAGAAACAGAGAGAAAGCTGTTTATTAACGCGAGCTTACCATCGAAGTC
OY 661 AAGGTGAGCGCGGACATGTCTCATTAACCGGTGGGTGAGGCTTGCAATCATGTGACG
DB 661 AAGGTGAGCGCGGACATGTCTCATTAACCGGTGGGTGAGGCTTGCAATCATGTGACG
OY 721 GAGATATTTCTCTGCAACAAGACAGAGGAATGCCGAGCCGGGTGGGGAAGATCAAC
DB 721 GAGATATTTCTCTGCAACAAGACAGAGGAATGCCGAGCCGGGTGGGGAAGATCAAC
OY 781 GATTACACACAGTGAACACCTTGTGTAAGTTGCAATACGCGCAATTTATTTGCTACAA
DB 781 GATTACACACAGTGAACACCTTGTGTAAGTTGCAATACGCGCAATTTATTTGCTACAA
OY 841 CGCAGCGCAGAGTGTGCGCGGACCGCGCCGCGGATTTGATTTGATCATGTCACG
DB 841 CGCAGCGCAGAGTGTGCGCGGACCGCGCCGCGGATTTGATTTGATCATGTCACG

```

```

DB 841 CGCAGCGCAGAGTGTGCGCGGACCGCGCCGCGGATTTGATTTGATCATGTCACG 900
OY 901 TTGACGCCCCATCCACCCGAAAAACAGGCGTATGGTGTGACATTAACCCACTCAGTACTG
DB 901 TTGACGCCCCATCCACCCGAAAAACAGGCGTATGGTGTGACATTAACCCACTCAGTACTG
OY 961 TTTATTTGCGGACAGCATACTAATCTGCGAAATCTCGGGGGCGCACTGAGCTCAACTGG
DB 961 TTTATTTGCGGACAGCATACTAATCTGCGAAATCTCGGGGGCGCACTGAGCTCAACTGG
OY 1021 ACGCTTCCCGGTCCAGCCGATTAACACGCGCCGAGGTGTGAGACTGTTTAAAGCTGG
DB 1021 ACGCTTCCCGGTCCAGCCGATTAACACGCGCCGAGGTGTGAGACTGTTTAAAGCTGG
OY 1081 CGTCGGCTTAAGCGATTAACAGCCAGTGTGATTCAGTGTGCTGCTCCAGACTTTACAG
DB 1081 CGTCGGCTTAAGCGATTAACAGCCAGTGTGATTCAGTGTGCTGCTCCAGACTTTACAG
OY 1141 CAGATGCGGTATTAACAGCCGCTGTCAATTAATACCGCGCCGAGAGGTGAACCTGACC
DB 1141 CAGATGCGGTATTAACAGCCGCTGTCAATTAATACCGCGCCGAGAGGTGAACCTGACC
OY 1201 CTGGCAGAGTGTGAAGAGCGGAATGCGAGGCGCATGTGTTGTTGGCAGGTTTACGGA
DB 1201 CTGGCAGAGTGTGAAGAGCGGAATGCGAGGCGCATGTGTTGTTGGCAGGTTTACGGA
OY 1261 ATCGTGAATGAAGCAGCATACCGGCTGCACTTGTGAGATCTCATACCATCACCATCAC
DB 1261 ATCGTGAATGAAGCAGCATACCGGCTGCACTTGTGAGATCTCATACCATCACCATCAC
OY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 2
AAD36473
ID AAD36473 standard; DNA; 1323 BP.
XX
AC AAD36473;
XX
DT 21-AUG-2002 (first entry)
XX
DE Escherichia coli phytase DNA.
XX
KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
  foodstuff; digestion; phytase; enzyme; gene; ds.
XX
OS Escherichia coli.
XX
FH Key
FT CDS 1..1323
FT FT 1..1323
FT FT /tag= a
FT FT /product= "E. coli phytase protein"
XX
PN WO200189317-A2.
XX
PD 29-NOV-2001.
XX
PF 15-MAY-2001; 2001WO-US15764.
XX
PR 25-MAY-2000; 2000US-0580937.
XX
PA (DIVE-) DIVERSA CORP.
XX
PI Short JM, Kretz KA, O'Donoghue E.
XX
XX WPI: 2002-164149/21.
XX DR P-PSDB: AAE22836.
XX
PT New dietary aids comprising sustained release biocompatible
  compositions, comprise agent that assists in digestion, useful for
  delivering enzymes, therapeutics, medicine or agents to an organism
  PT

```

XX Disclosure; Fig 1; 89pp; English.

PS The present invention relates to novel dietary aids comprising sustained
CC release biocompatible composition which comprises an agent (enzymes such
CC as phytase, amylase, esterase, protease) that assists in digestion. The
CC biocompatible composition is effective upon oral consumption and release
CC in the digestive tract of a subject. The dietary aids are useful for
CC delivering enzymes, therapeutics, medicine and agents to an organism.
CC The use of enzymes and other agents in digestive aids of livestock or
CC domesticated animals not only improves the animal's health and life
CC expectancy but also assists in increasing the health of livestock or
CC in the production of foodstuffs from livestock. The present sequence
CC is Escherichia coli phytase DNA.

XX Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 1 other;

Query Match 99.9%; Score 1322; DB 24; Length 1323;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCGTTAACCCCGCAATGTGA 60
DB 1 ATGAAAGCGATCTTAATCCATTTTATCTCTCTGATTCGTTAACCCCGCAATGTGA 60
QY 61 TTGGCTCAGAGTGAAGCGGAGTGAAGCTGAAGCTGAGTGTGATGTCATGCTCATGT 120
DB 61 TTGGCTCAGAGTGAAGCGGAGTGAAGCTGAAGCTGAGTGTGATGTCATGCTCATGT 120
QY 121 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGATGATGTCACCCGAGAGCATGSCA 180
DB 121 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGATGATGTCACCCGAGAGCATGSCA 180
QY 121 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGATGATGTCACCCGAGAGCATGSCA 180
DB 121 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGATGATGTCACCCGAGAGCATGSCA 180
QY 181 ACCTGCGCGGTAATCTGGTGTGCTGACACCGGAGGCTGAGTGAATGCTTATCTC 240
DB 181 ACCTGCGCGGTAATCTGGTGTGCTGACACCGGAGGCTGAGTGAATGCTTATCTC 240
QY 181 ACCTGCGCGGTAATCTGGTGTGCTGACACCGGAGGCTGAGTGAATGCTTATCTC 240
DB 181 ACCTGCGCGGTAATCTGGTGTGCTGACACCGGAGGCTGAGTGAATGCTTATCTC 240
QY 241 GGACATTACCAAGCCAGCGCTGTGTAGCCGAGGAGTGTGCGGAGAAAGGCTGCGG 300
DB 241 GGACATTACCAAGCCAGCGCTGTGTAGCCGAGGAGTGTGCGGAGAAAGGCTGCGG 300
QY 301 CAGTCTGTCTCAGTGTGCGATTAATGCTGATGTGACAGAGGTCACCCGTAACAGAGCAA 360
DB 301 CAGTCTGTCTCAGTGTGCGATTAATGCTGATGTGACAGAGGTCACCCGTAACAGAGCAA 360
QY 301 CAGTCTGTCTCAGTGTGCGATTAATGCTGATGTGACAGAGGTCACCCGTAACAGAGCAA 360
DB 301 CAGTCTGTCTCAGTGTGCGATTAATGCTGATGTGACAGAGGTCACCCGTAACAGAGCAA 360
QY 361 GCGTGTGCGCGCGGCGCTGCGACCTGACTGTGCAATAACCGTACATACCCAGGAGATAGC 420
DB 361 GCGTGTGCGCGCGGCGCTGCGACCTGACTGTGCAATAACCGTACATACCCAGGAGATAGC 420
QY 421 TTGAGTCCGAGTCCGTTATTAATCTCTTAATAAAGCTGCTTGGCCAACTGAGTAACGCG 480
DB 421 TTGAGTCCGAGTCCGTTATTAATCTCTTAATAAAGCTGCTTGGCCAACTGAGTAACGCG 480
QY 481 AAGGTACTGACGATCCTCAGCAGAGGAGGTCATTTGCTGACTTTACCGGGCAT 540
DB 481 AAGGTACTGACGATCCTCAGCAGAGGAGGTCATTTGCTGACTTTACCGGGCAT 540
QY 541 CGGCAAAAGCGGCTTTGGCAACTGGAAGGCTGCTTAATTTTCGCAATCAACTGTGCG 600
DB 541 CGGCAAAAGCGGCTTTGGCAACTGGAAGGCTGCTTAATTTTCGCAATCAACTGTGCG 600
QY 601 CTTAAAGGTGAGAAAGAGGAGGAGGAGTTCATTAACGAGCATTAACCATGGAATC 660
DB 601 CTTAAAGGTGAGAAAGAGGAGGAGGAGTTCATTAACGAGCATTAACCATGGAATC 660
QY 661 AAGGTAGCGCCGACATGTCTCATTAACCGGTGCGGTAGAGCTGCGATCAATGCTGAGC 720
DB 661 AAGGTAGCGCCGACATGTCTCATTAACCGGTGCGGTAGAGCTGCGATCAATGCTGAGC 720
QY 721 GAGATATTTCTCTGCAACAGCAAGGAGTCCGAGGCGGAGGTGGGAAGATCACC 780
DB 721 GAGATATTTCTCTGCAACAGCAAGGAGTCCGAGGCGGAGGTGGGAAGATCACC 780

QY 781 GATTACACAGTGAAGAACCTTGTAGTTGATTAACCGCAATTTATTTGCTACAA 840
DB 781 GATTACACAGTGAAGAACCTTGTAGTTGATTAACCGCAATTTATTTGCTACAA 840
QY 841 CGCAGCGGAGGTTGCGCGGAGCGGCGCAACCCGTTATTTGATTAATCAGGAGCG 900
DB 841 CGCAGCGGAGGTTGCGCGGAGCGGCGCAACCCGTTATTTGATTAATCAGGAGCG 900
QY 901 TTGAGCGGCGGATCCAGCGCAAAAGAGGCGGATGATGATTAACCATCTGATCTG 960
DB 901 TTGAGCGGCGGATCCAGCGCAAAAGAGGCGGATGATGATTAACCATCTGATCTG 960
QY 961 TTTATTTCCGCGACATGATTAATCTGCAAAATCTCGGCGGCGGAGTGAAGTCTG 1020
DB 961 TTTATTTCCGCGACATGATTAATCTGCAAAATCTCGGCGGCGGAGTGAAGTCTG 1020
QY 1021 ACCTTCCGCGGATGCGGAGTGAAGCAGCGCGGAGTGAAGTCTGTTGAAGCTG 1080
DB 1021 ACCTTCCGCGGATGCGGAGTGAAGCAGCGCGGAGTGAAGTCTGTTGAAGCTG 1080
QY 1081 CGTGGCTAAGCGATTAACAGCGAGTGTGATTTGCTGCTTCCAGACTTTACAG 1140
DB 1081 CGTGGCTAAGCGATTAACAGCGAGTGTGATTTGCTGCTTCCAGACTTTACAG 1140
QY 1141 CAGATGCGTGTAAAGCGCGCTGTCAATTAATACGCGCGGAGAGGTAAGTCAAC 1200
DB 1141 CAGATGCGTGTAAAGCGCGCTGTCAATTAATACGCGCGGAGAGGTAAGTCAAC 1200
QY 1201 CTGCGAGGATGTAAGAGCGAATATGCGAGGCGATGTTGCTGAGGAGGTTTACGCA 1260
DB 1201 CTGCGAGGATGTAAGAGCGAATATGCGAGGCGATGTTGCTGAGGAGGTTTACGCA 1260
QY 1261 ATCTGATTAAGAGCAGCATACCGGCGGCTGATGATGATCTACATCAGATCAG 1320
DB 1261 ATCTGATTAAGAGCAGCATACCGGCGGCTGATGATGATCTACATCAGATCAG 1320
QY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 3
AAD25460
ID AAD25460 standard; DNA; 1323 BP.
XX AAD25460;
AC 26-MAR-2002 (first entry)
XX
DE Escherichia coli B phytase DNA.
KW Bacterial phytase; K12 appa phytase; protease stability; anabolic;
KW gastrointestinal; nutritional value; feed treatment process; therapy;
KW thermal tolerance; growth performance; alcoholic drink; biopulping;
KW non-alcoholic drink; biobleaching; B phytase; ds.
OS Escherichia coli.
XX
XX
FH Key location/Qualifiers
FT 1..1323
FT CDS /tag- a
FT /product= "E. coli B phytase protein"
PN W0200190333-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001MO-US17118.
XX
XX 25-MAY-2000; 2000US-0580515.
XX
XX (DIVE-) DIVERSA CORP.
XX

PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;
 XX WPI: 2002-083108/11.
 DR P-PSDB: AEI15806.
 XX
 PT New bacterial phytase for e.g. improving the nutritional value of
 PT phytase-containing foodstuffs and subsequently improving the growth
 PT performance of an organism that consumes it, or in treating animal
 PT digestive systems
 PS
 PS Claim 2; Fig 1; 170pp; English.

XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
 CC K12 appa phytase. The enzyme has phytase activity and improved protease
 CC tolerance when compared with wild-type phytase. It has improved thermal
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytase-containing foodstuffs and subsequently
 CC improving the growth performance of an organism that consumes it, in
 CC treating animal digestive systems, in feed treatment processes and for
 CC in vitro purposes related to research, discovery and development. They
 CC are also used for generating recombinant digestive system life forms,
 CC for producing or manufacturing alcoholic and non-alcoholic drinks based
 CC on the use of moulds, grains and/or plants, in biopulping and bio-
 CC bleaching where a reduction in the use of environmentally harmful
 CC chemicals that are traditionally used in the pulp and paper industry
 CC is desired and in the reduction or possible elimination of the need
 CC for mineral supplements, enzymes or therapeutic drugs for animals
 CC from the daily feed thus increasing the amount calories and nutrients
 CC present in the feed. The present sequence is a DNA encoding *E. coli*
 CC B phytase protein.

XX Sequence 1323 BP: 323 A; 353 C; 357 G; 289 T; 1 other:

Query Match 99.9%; Score 1322; DB 24; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCTGCAATCTGCA 60
 DB 1 ATGAAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCTGCAATCTGCA 60
 QY 61 TTGCTCGAGTGGAGCCGAGCTGAAGTGGAAAGTGGTGGATTTGATGCTGATATGT 120
 DB 61 TTGCTCGAGTGGAGCCGAGCTGAAGTGGAAAGTGGTGGATTTGATGCTGATATGT 120
 QY 121 GTGCGTGTCCACCAAGGCGCAGCAACTGATGAGATGTACCCAGAGCATGGCA 180
 DB 121 GTGCGTGTCCACCAAGGCGCAGCAACTGATGAGATGTACCCAGAGCATGGCA 180
 QY 121 GTGCGTGTCCACCAAGGCGCAGCAACTGATGAGATGTACCCAGAGCATGGCA 180
 DB 121 GTGCGTGTCCACCAAGGCGCAGCAACTGATGAGATGTACCCAGAGCATGGCA 180
 QY 181 ACCTGGCCGTTAAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
 DB 181 ACCTGGCCGTTAAACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 240
 QY 241 GGACATTAACCAAGCCGAGCTGTGAGCCGAGATGTGGGCAAAAAGGCTGCCG 300
 DB 241 GGACATTAACCAAGCCGAGCTGTGAGCCGAGATGTGGGCAAAAAGGCTGCCG 300
 QY 301 CAGTCTGTGAGTGGAGTATTTGATGATGTCGACGAGCTTACCCGTAACAGAGCGAA 360
 DB 301 CAGTCTGTGAGTGGAGTATTTGATGATGTCGACGAGCTTACCCGTAACAGAGCGAA 360
 QY 301 CAGTCTGTGAGTGGAGTATTTGATGATGTCGACGAGCTTACCCGTAACAGAGCGAA 360
 DB 301 CAGTCTGTGAGTGGAGTATTTGATGATGTCGACGAGCTTACCCGTAACAGAGCGAA 360
 QY 361 GCTTTCGCGCGGCGGCTGGACCTGACTGTGCAATACCGTAATACCGAGAGATAG 420
 DB 361 GCTTTCGCGCGGCGGCTGGACCTGACTGTGCAATACCGTAATACCGAGAGATAG 420
 QY 421 TCCAGTCCCGATCCGTTATTTAATCCCTAAAACTGGCTTTGCAACTGGATTAACCG 480
 DB 421 TCCAGTCCCGATCCGTTATTTAATCCCTAAAACTGGCTTTGCAACTGGATTAACCG 480
 QY 481 AACGTGACTGACGAGTCTCAGCAGGAGGAGGATGCAATTTGCTGACTTACCGGCGAT 540
 DB 481 AACGTGACTGACGAGTCTCAGCAGGAGGAGGATGCAATTTGCTGACTTACCGGCGAT 540

QY 541 CGGCAAAAGCGGCTTTCGCAACGAGGAGGCTTAATTTTCCGCAATCAACTGTCG 600
 DB 541 CGGCAAAAGCGGCTTTCGCAACGAGGAGGCTTAATTTTCCGCAATCAACTGTCG 600
 QY 601 CTTAAACGTGAGAAACAGAGCAAGCTGTTTATTACCGAGGCTTACCATGGAATC 660
 DB 601 CTTAAACGTGAGAAACAGAGCAAGCTGTTTATTACCGAGGCTTACCATGGAATC 660
 QY 661 AAGGTGAGCGCGGAGCAATGTCTCATTACCGGTGGGATGAGCTTCATGCTGACG 720
 DB 661 AAGGTGAGCGCGGAGCAATGTCTCATTACCGGTGGGATGAGCTTCATGCTGACG 720
 QY 721 GAGATATTTTCCTGCAACAGCAAGGAGGATCCGAGCGCGGCTGGGAGGATGAC 780
 DB 721 GAGATATTTTCCTGCAACAGCAAGGAGGATCCGAGCGCGGCTGGGAGGATGAC 780
 QY 781 GATTACACCAAGTGGAAACCTGTTGATGTTGATTAACCGGCAATTTATTTGTTAA 840
 DB 781 GATTACACCAAGTGGAAACCTGTTGATGTTGATTAACCGGCAATTTATTTGTTAA 840
 QY 841 CGCAGCGCAGAGTGGCGCGAGCGCGCAGCGCGCGGATTTGATGATGATGACG 900
 DB 841 CGCAGCGCAGAGTGGCGCGAGCGCGCAGCGCGCGGATTTGATGATGATGACG 900
 QY 901 TTGACGCGCCATCCACCGCAAAAACAGCGGATGTTGATGATTAACCATTCAGT 960
 DB 901 TTGACGCGCCATCCACCGCAAAAACAGCGGATGTTGATGATTAACCATTCAGT 960
 QY 961 TTTATGCGCGGACAGTACTAATGTTGCAATGCGCGCGGAGCTGAGTCAACTG 1020
 DB 961 TTTATGCGCGGACAGTACTAATGTTGCAATGCGCGCGGAGCTGAGTCAACTG 1020
 QY 1021 ACGCTTCCGCGTACCGGATTAACAGCGCGCGCGGAGTGTGATGATGTTGAACG 1080
 DB 1021 ACGCTTCCGCGTACCGGATTAACAGCGCGCGCGGAGTGTGATGATGTTGAACG 1080
 QY 1081 CGTGGCGTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1140
 DB 1081 CGTGGCGTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1140
 QY 1141 CAGATGCGTGAATTAACCGCGCTGTCATTAATATACCGCGCGGAGAGTGAAC 1200
 DB 1141 CAGATGCGTGAATTAACCGCGCTGTCATTAATATACCGCGCGGAGAGTGAAC 1200
 QY 1201 CTGGCAGAGTGTGAAGAGCGGAATGCGCAGGCGATGTGCTGGTGGAGTTTAC 1260
 DB 1201 CTGGCAGAGTGTGAAGAGCGGAATGCGCAGGCGATGTGCTGGTGGAGTTTAC 1260
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGTCATTAATATACCGCATCACAT 1320
 DB 1261 ATCGTGAATGAAGCAGCATACCGCGCTGTCATTAATATACCGCATCACAT 1320
 QY 1321 TAA 1323
 DB 1321 TAA 1323

RESULT 4
 AAD25463
 ID AAD25463 standard: DNA; 1901 BP.
 AAD25463;
 26-MAR-2002 (first entry)
 Escherichia coli appa phytase wild type DNA.
 Bacterial phytase; K12 appa phytase; protease stability; anabolic;
 gastrointestinal; nutritional value; feed treatment process; therapy;
 thermal tolerance; growth performance; alcoholic drink; biopulping;
 non-alcoholic drink; biobleaching; ds.
 Escherichia coli.

XX Key Location/Qualifiers
FH CDS 188..1486
FT /tag= a
FT /Product= "E. coli appa phytase protein"
PN M0200190333-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001MO-US17118.
XX
XX 25-MAY-2000; 2000US-0580515.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;
XX P-PSDB; AAE15807.
XX
XX MPI; 2002-083108/11.
XX
XX P-PSDB; AAE15807.
XX
XX New bacterial phytase for e.g. improving the nutritional value of
PT phytase-containing foodstuffs and subsequently improving the growth
PT performance of an organism that consumes it, or in treating animal
PT digestive systems
XX
XX Claim 42: Fig 7; 170pp; English.
XX
XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
CC K12 appa phytase. The enzyme has phytase activity and improved thermal
CC tolerance when compared with wild-type phytase. It has improved protease
CC stability at low pH. The recombinant phytase is useful for improving the
CC nutritional value of phytase-containing foodstuffs and subsequently
CC improving the growth performance of an organism that consumes it, in
CC treating animal digestive systems, in feed treatment processes and for
CC in vitro purposes related to research, discovery and development. They
CC are also used for generating recombinant digestive system life forms,
CC for producing or manufacturing alcoholic and non-alcoholic drinks based
CC on the use of moulds, grains and/or plants, in bioprocessing and bio-
CC bleaching where a reduction in the use of environmentally harmful
CC chemicals that are traditionally used in the pulp and paper industry
CC is desired and in the reduction or possible elimination of the need
CC for mineral supplements, enzymes or therapeutic drugs for animals
CC from the daily feed thus increasing the amount calories and nutrients
CC present in the feed. The present sequence is a DNA encoding E. coli
CC appa phytase wild type protein.
XX
SQ Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 1 other;
Query Match 97.3%; Score 1287.8; DB 24; Length 1901;
Best local similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCCGCAATCTCA 60
DB 188 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCCGCAATCTCA 247
OY 61 TTGCGTCAGAGTAGCGGAGCTGGAAGCTGGAAGTGTGATGTCATCTGTCATGCT 120
DB 248 TTGCGTCAGAGTAGCGGAGCTGGAAGCTGGAAGTGTGATGTCATCTGTCATGCT 307
OY 121 GTGCGTGTCTCAACAGGCGGAGCTGGAAGCTGGAAGTGTGATGTCATCTGTCATGCT 180
DB 308 GTGCGTGTCTCAACAGGCGGAGCTGGAAGCTGGAAGTGTGATGTCATCTGTCATGCT 367
OY 181 ACCTGCGCGGTAAACTGCGGTGCTGACACCGCGGCTGCTGAGCTAATCGCTATCTC 240
DB 368 ACCTGCGCGGTAAACTGCGGTGCTGACACCGCGGCTGCTGAGCTAATCGCTATCTC 427
OY 241 GGCATTACCAAGCGGAGCTGCTGTCAGCGAGGATGCTGGCGAAAAAGGCTGCGCG 300
DB 428 GGCATTACCAAGCGGAGCTGCTGTCAGCGAGGATGCTGGCGAAAAAGGCTGCGCG 487

OY 301 CAGCTGTGTCAGGTCGCGATTAATGCTGATGTCAGAGCGGTACCCGTAACAGCGGAA 360
DB 488 CAGCTGTGTCAGGTCGCGATTAATGCTGATGTCAGAGCGGTACCCGTAACAGCGGAA 547
OY 361 GCGTTGCGCGCGGCGGTGCGACCTGACTGTGCAATTAACGTCATACCAAGCAGATACG 420
DB 548 GCGTTGCGCGCGGCGGTGCGACCTGACTGTGCAATTAACGTCATACCAAGCAGATACG 607
OY 421 TCCAGTCCCGATCCGTTATTTAATCCCTTAATAAAGCGCTTTGCCAATCGATTAACGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTAATCCCTTAATAAAGCGCTTTGCCAATCGATTAACGCG 667
OY 481 AACGTGACGTACGCGATCCTCAGAGGCGAGAGGCTCAATTCGTGACTTACCGGCGAT 540
DB 668 AACGTGACGTACGCGATCCTCAGAGGCGAGAGGCTCAATTCGTGACTTACCGGCGAT 727
OY 541 CGGCAAAAGCGGCTTTCGCGAATGGAACGGGTCTTAATTTCCGCAATCAACTGTGCG 600
DB 728 CGGCAAAAGCGGCTTTCGCGAATGGAACGGGTCTTAATTTCCGCAATCAACTGTGCG 787
OY 601 CTTAAAGGTGAGAAACAGAGCAAGAGCTGTCATTAACGAGCATTAACATGGAATC 660
DB 788 CTTAAAGGTGAGAAACAGAGCAAGAGCTGTCATTAACGAGCATTAACATGGAATC 847
OY 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGGTAGCCCTGCATCAATGCTGACG 720
DB 848 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGGTAGCCCTGCATCAATGCTGACG 907
OY 721 GAGATATTTCTCTGCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 780
DB 908 GAGATATTTCTCTGCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 967
OY 781 GATTCAACAGAGGAAACACCTGCTTAAGTTGTCATTAACCGGCAATTTTTCGTCACAA 840
DB 968 GATTCAACAGAGGAAACACCTGCTTAAGTTGTCATTAACCGGCAATTTTTCGTCACAA 1027
OY 841 CGCAGCGCAAGGTTGCG 900
DB 1028 CGCAGCGCAAGGTTGCG 1087
OY 901 TTGACGCGCGATCCACCGCAACAGCGGCTATGTCATTAACCACTTACGATCTG 960
DB 1088 TTGACGCGCGATCCACCGCAACAGCGGCTATGTCATTAACCACTTACGATCTG 1147
OY 961 TTATGCGCGGACAGCATTAATCTGGAATCTGCGGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 1148 TTATGCGCGGACAGCATTAATCTGGAATCTGCGGCGCGCGCGCGCGCGCGCGCGCG 1207
OY 1021 ACCTTCCCGGTGACGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1208 ACCTTCCCGGTGACGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1267
OY 1081 CGTGGCTAAGCGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1268 CGTGGCTAAGCGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1327
OY 1141 CAGATGCGGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
DB 1328 CAGATGCGGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1387
OY 1201 CTGCGAGGATTAAGAGCGGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCG 1260
DB 1388 CTGCGAGGATTAAGAGCGGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCG 1447
OY 1261 ATCGTGAATGAAGCAGCATACCGGCGGTGCAAGTTGAGAT 1300
DB 1448 ATCGTGAATGAAGCAGCATACCGGCGGTGCAAGTTGAGAT 1487
RESULT 5
ABK12514
ID ABK12514 standard; DNA; 1299 BP.
XX

AC ABK12514;
 XX 05-JUN-2002 (first entry)
 XX DNA encoding phytase associated protein.
 DE phytase; gene; ds.
 XX unidentified.
 OS
 XX
 FT Key location/Qualifiers
 FT CDS 1..1299
 FT /*tag= a
 FT /product= "Phytase associated protein"
 XX KR9086028-A.
 XX PD 15-DEC-1999.
 XX PF 25-MAY-1998; 98KR-0018810.
 XX PR 25-MAY-1998; 98KR-0018810.
 XX PA (MOOI-) MOOJIN CO LTD.
 XX PI Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 XX WPI: 2000-645078/62.
 XX DR P-PSDB: AAU77775.
 XX PT Novel phytase gene, recombinant phytase and usage thereof -
 XX PS Claim 1; Fig 2; 10pp; Korean.
 XX CC The invention relates to a novel phytase gene, a recombinant
 CC phytase gene and their uses. This sequence encodes a phytase
 CC associated protein, described in the invention.
 CC XX
 SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 other;
 Query Match 97.3%; Score 1287; DB 21; Length 1299;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1290; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 421 TCCAGTCCCGATCCGTTATTTAACTCTTAAAACTGGCCGTTTCCCAACTGGATTAACGCG 480
 481 AAGTGACTGACGCGATTCCTCAGCAGGCGACAGAGGTCATTTGCTATTACGGGCAT 540
 481 AAGTGACTGACGCGATTCCTCAGCAGGCGACAGAGGTCATTTGCTATTACGGGCAT 540
 541 CGGCAAAACGGCGTTTCGGCACTGAAACGGGTCCTTAATTTCCGCAATCAACTGTCG 600
 541 CGGCAAAACGGCGTTTCGGCACTGAAACGGGTCCTTAATTTCCGCAATCAACTGTCG 600
 601 CTTAAAGGTGAGAAACAGAGAGAAAGCTGTTCAATTAACGCAAGCATTAACATCGGAATC 660
 601 CTTAAAGGTGAGAAACAGAGAGAAAGCTGTTCAATTAACGCAAGCATTAACATCGGAATC 660
 661 AAGGTGAGCGCGCAATGTCATTAACCGGTCGGGTAAAGCTTCGATCAATCTGACG 720
 661 AAGGTGAGCGCGCAATGTCATTAACCGGTCGGGTAAAGCTTCGATCAATCTGACG 720
 721 GAGATATTTCTCTGCAACAAACAGAGGAATGCCGAGCCGGGTCGGGAAGATCAAC 780
 721 GAGATATTTCTCTGCAACAAACAGAGGAATGCCGAGCCGGGTCGGGAAGATCAAC 780
 781 GATTCAACACAGTGGAAACACCTTGCTAAGTTTGATTAACGGCAATTTATTCTCTCAA 840
 781 GATTCAACACAGTGGAAACACCTTGCTAAGTTTGATTAACGGCAATTTATTCTCTCAA 840
 841 CGCAGCGCAGAGGTTGGCCCGCAGCCGCCACCCGTTATGATTAACAGACAGCG 900
 841 CGCAGCGCAGAGGTTGGCCCGCAGCCGCCACCCGTTATGATTAACAGACAGCG 900
 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGATGATTAACCGCAATTTATTCTCTCAA 960
 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGATGATTAACCGCAATTTATTCTCTCAA 960
 961 TTTATGCGCGGACACGATTAATCTGCAAAATCTCGCGGCGACGCTGAGACTCACTCG 1020
 961 TTTATGCGCGGACACGATTAATCTGCAAAATCTCGCGGCGACGCTGAGACTCACTCG 1020
 1021 ACGCTCCCGGTCAGCGGATTAACAGCGCGCGAGGTGTAAGTGTGTTGAACGCTG 1080
 1021 ACGCTCCCGGTCAGCGGATTAACAGCGCGCGAGGTGTAAGTGTGTTGAACGCTG 1080
 1081 CGTGGGCTAACGATTAACAGCGCGAGGTGTAAGTGTGTTGAACGCTG 1140
 1081 CGTGGGCTAACGATTAACAGCGCGAGGTGTAAGTGTGTTGAACGCTG 1140
 1141 CAGATGCGGTGTAAGAGCGCGCTGCTAATTAATACGCGCGCGAGAGGTAAACTGACC 1200
 1141 CAGATGCGGTGTAAGAGCGCGCTGCTAATTAATACGCGCGCGAGAGGTAAACTGACC 1200
 1201 CTGGCAGATGTAAGAGCGCGAGGTGTAAGTGTGTTGAACGCTG 1260
 1201 CTGGCAGATGTAAGAGCGCGAGGTGTAAGTGTGTTGAACGCTG 1260
 1261 ATCGTGAATGAAGCAGCATACCGGCTGCACTTTG 1296
 1261 ATCGTGAATGAAGCAGCATACCGGCTGCACTTTG 1296

RESULT 6
 AAC68298 standard; DNA; 3470 BP.
 AAC68298:
 20-FEB-2001 (first entry)
 R15/APPA plasmid coding sequence.
 Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 environmental pollution; pig; ds.
 XX

OS Rattus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX PN W0200064247-A1.
 XX PD 02-NOV-2000.
 XX PF 20-APR-2000; 2000MO-CA00430.
 XX PR 23-APR-1999; 99US-0130508.
 XX PA (UYGU-) UNIV GUELPH.
 PI Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 DR P-PDB; AAB36261.
 XX PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX PS Claim 14; Fig 21; 152pp; English.
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APFA phytase coding sequence.
 SO Sequence 3470 BP; 1065 A; 721 C; 735 G; 949 T; 0 other;

Query Match 96.98; Score 1282.2; DB 21; Length 3470;
 Best Local Similarity 99.38; Pred. No. 0;
 Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1. ATGAAGCGATCTTATCCATTTTATCTCTGATCCGTTAACCCCGCAATCGCA 60
 DB 1811 ATGAAGCGATCTTATCCATTTTATCTCTGATCCGTTAACCCCGCAATCGCA 1870
 QY 61 TTGCGCTCAGAGTGGAGCGGAGTGAAGTGAAGTGGTGAATTCAGTCTCATGT 120
 DB 1871 TTGCGCTCAGAGTGGAGCGGAGTGAAGTGAAGTGGTGAATTCAGTCTCATGT 1930
 QY 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGAGCATG 180
 DB 1931 GTGCGTGTCTCAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGAGCATG 1990
 QY 181 ACCTGCGCGGTAAACTGGGTGTGACACCGCGGAGTGTGATGCTATTCGCTATCTC 240
 DB 1991 ACCTGCGCGGTAAACTGGGTGTGACACCGCGGAGTGTGATGCTATTCGCTATCTC 2050
 QY 241 GGACATTACCAAGCGGAGTGTGATGACCGGAGTGTGATGCTATTCGCTATCTC 300
 DB 2051 GGACATTACCAAGCGGAGTGTGATGACCGGAGTGTGATGCTATTCGCTATCTC 2110
 QY 301 CAGTGTGTGATGCTATTCGCTATGATGTCGACAGAGGTACCCGTTAAACAGGCGAA 360
 DB 2111 CAGTGTGTGATGCTATTCGCTATGATGTCGACAGAGGTACCCGTTAAACAGGCGAA 2170
 QY 361 GCGTTGGCGCGGCGGCTGGACCTGACGTGCAATAACCGTACATACCCAGGAGATACG 420
 DB 2171 GCGTTGGCGCGGCGGCTGGACCTGACGTGCAATAACCGTACATACCCAGGAGATACG 2230
 QY 421 TCCAGTCCCGATCGTTATTTATCTCTTAAATAAGTGGGTTGCCAAGTGAATACGCG 480
 DB 2231 TCCAGTCCCGATCGTTATTTATCTCTTAAATAAGTGGGTTGCCAAGTGAATACGCG 2290
 QY 481 AACGTGACTGACGCGATCTCTACAGAGGCAAGAGGCTCAATTCGTAACCTTACGGGCAT 540
 DB 2291 AACGTGACTGACGCGATCTCTCTACAGAGGCAAGAGGCTCAATTCGTAACCTTACGGGCAT 2350

QY 541 CGGCAAAAGCGGCTTTGCGGCAACGTAAGCGGCTGCTAATTTCCGAATCAACTGTGC 600
 DB 2351 CGGCAAAAGCGGCTTTGCGGCAACGTAAGCGGCTGCTAATTTCCGAATCAACTGTGC 2410
 QY 601 CTTAAACGTCGAGAAACAGAGCAAGCAAGCTGTTCAATTAACGAGCATTTACATCGAATC 660
 DB 2411 CTTAAACGTCGAGAAACAGAGCAAGCAAGCTGTTCAATTAACGAGCATTTACATCGAATC 2470
 QY 661 AAGGTAGCGCGGCAACATGTCATTTAACCGGTGCGGTAAAGCTCGCATCATGTCAGC 720
 DB 2471 AAGGTAGCGCGGCAACATGTCATTTAACCGGTGCGGTAAAGCTCGCATCATGTCAGC 2530
 QY 721 GAGATATTTCTCTGCAACAGCAAGCAAGGGAATCCGAGCGGGGGGGAAGATATAC 780
 DB 2531 GAGATATTTCTCTGCAACAGCAAGCAAGGGAATCCGAGCGGGGGGGAAGATATAC 2590
 QY 781 GATTCAACACGATGTAACACCTTGTCTAAGTTGCTAATACCGCAATTTTATTTGCTACAA 840
 DB 2591 GATTCAACACGATGTAACACCTTGTCTAAGTTGCTAATACCGCAATTTTATTTGCTACAA 2650
 QY 841 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCGTTATTTGATTCATGACAGCG 900
 DB 2651 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCGTTATTTGATTCATGACAGAGCG 2710
 QY 901 TTGACGCGCCATCCACCGCAAAAACAGGCGTATGTTGATTCATCCACTTACATG 960
 DB 2711 TTGACGCGCCATCCACCGCAAAAACAGGCGTATGTTGATTCATCCACTTACATG 2770
 QY 961 TTTATTTGCGCGAGACGATTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTGG 1020
 DB 2771 TTTATTTGCGCGAGACGATTAATCTGCAAAATCTGCGCGGCGCACTGAGCTCAACTGG 2830
 QY 1021 ACCTTCCCGGTGACCGGATTAACAGCGCGCCAGGTGTAATCTGTTTGAACGCTGG 1080
 DB 2831 ACCTTCCCGGTGACCGGATTAACAGCGCGCCAGGTGTAATCTGTTTGAACGCTGG 2890
 QY 1081 CGTGGCGTAAGCATTAACAGCGAGTGTGATTCAGGTTTCCGTGCTTCCAGACTTTACAG 1140
 DB 2891 CGTGGCGTAAGCATTAACAGCGAGTGTGATTCAGGTTTCCGTGCTTCCAGACTTTACAG 2950
 QY 1141 CAGATGCGTATTAACAGCGCGCTGTCAATTAATACCGCGCGGAGAGGTGAATCTGAC 1200
 DB 2951 CAGATGCGTATTAACAGCGCGCTGTCAATTAATACCGCGCGGAGAGGTGAATCTGAC 3010
 QY 1201 CTGGCAGAGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGGAGTTTACGCAA 1260
 DB 3011 CTGGCAGAGATGTAAGAGCGAAATGCGCAGGCGATGTTGTTGGAGTTTACGCAA 3070
 QY 1261 ATGCTGAATGAACGACGATACCGGCGGTGAGTTTG 1296
 DB 3071 ATGCTGAATGAACGACGATACCGGCGGTGAGTTTG 3106

RESULT 7
 AAC68296
 ID AAC68296 standard; DNA: 4060 BP.
 XX AAC68296;
 DE 20-FEB-2001 (first entry)
 XX
 DE R15/APFA plasmid coding sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig; ds.
 XX
 OS Rattus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX PN W0200064247-A1.
 PD 02-NOV-2000.

XX 20-APR-2000; 2000MO-CA00430.
 PF
 XX 23-APR-1999; 99US-0130508.
 PR
 XX (UYGU-) UNIV GUELPH.
 PA
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI: 2000-687245/67.
 DR P-PSDB: AAB36259.
 XX
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein -
 PS Claim 14; Fig 19; 152pp; English.
 CC
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.
 CC
 SQ Sequence 4060 BP; 1257 A; 814 C; 843 G; 1146 T; 0 other;
 Query Match 96.9%; Score 1282.2; DB 21; Length 4060;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 ATGAAGCGATCTTATCCCATTTTATCTCTCTGATTCGTTACCCCGCAATCTGCA 60
 DB 1811 ATGAAGCGATCTTATCCCATTTTATCTCTCTGATTCGTTACCCCGCAATCTGCA 1870
 QY 61 TTGCTGAGAGTGAGCGGAGCTGAGCTGAGAAAGTGTGATTCAGTTCATGCT 120
 DB 1871 TTGCTGAGAGTGAGCGGAGCTGAGCTGAGAAAGTGTGATTCAGTTCATGCT 1930
 QY 121 GTGCTGCTCCAAACCAAGCAGCAGCACTGATGAGAGATGACCCAGAGAGATGGCA 180
 DB 1931 GTGCTGCTCCAAACCAAGCAGCAGCACTGATGAGAGATGACCCAGAGAGATGGCA 1990
 QY 181 ACCTGGCCGTTAAACCTGGTGTGACACACCGCGAGTGTGAGTAACTGCCTATCTC 240
 DB 1991 ACCTGGCCGTTAAACCTGGTGTGACACACCGCGAGTGTGAGTAACTGCCTATCTC 2050
 QY 241 GGNATTAACCAACCGCCGCTGTGACCGCAGCGATTCGCGGGAAGAGGCTGCCG 300
 DB 2051 GGNATTAACCAACCGCCGCTGTGACCGCAGCGATTCGCGGGAAGAGGCTGCCG 2110
 QY 301 CAGTCTGTGAGTGCGGATTAATGCTGATGTGACAGAGCGTACCCGTAACAGGCGAA 360
 DB 2111 CAGTCTGTGAGTGCGGATTAATGCTGATGTGACAGAGCGTACCCGTAACAGGCGAA 2170
 QY 421 TTCAGTCCGATCCGTTATTTAAATCCTTAAACCTGGCGTTCCCAACAGGAGGCG 480
 DB 2231 TTCAGTCCGATCCGTTATTTAAATCCTTAAACCTGGCGTTCCCAACAGGAGGCG 2290
 QY 481 AACGTGACGACGATCTCAGCAGGAGGAGGAGGATTCGTTGACTTACCGGCGAT 540
 DB 2291 AACGTGACGACGATCTCAGCAGGAGGAGGAGGATTCGTTGACTTACCGGCGAT 2350
 QY 541 CGGCAAAAGCGGTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGGC 600
 DB 2351 CGGCAAAAGCGGTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGGC 2410
 QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGCAAGGATTAACATCGGAATC 660

DB 2411 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGAGGATTAACATCGGAATC 2470
 QY 661 AAGGTGAGCGCCGCAATGTCATTAACCGGCTGGTAAAGCCCGCATCAATCTGACG 720
 DB 2471 AAGGTGAGCGCCGCAATGTCATTAACCGGCTGGTAAAGCCCGCATCAATCTGACG 2530
 QY 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCGGGGTGGGAAGATCAAC 780
 DB 2531 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCGGGGTGGGAAGATCAAC 2590
 QY 781 GATTACACACAGTGGAAACACCTTCTTAAGTTTGCATTAACGCGCAATTTTATTTCTCA 840
 DB 2591 GATTACACACAGTGGAAACACCTTCTTAAGTTTGCATTAACGCGCAATTTTATTTCTCA 2650
 QY 841 CGCAGCGCAGAGGTTGCGCGGAGCGGCGCCACCCGTTATTTGATTTGATTCAGCAGCG 900
 DB 2651 CGCAGCGCAGAGGTTGCGCGGAGCGGCGCCACCCGTTATTTGATTTGATTCAGCAGCG 2710
 QY 901 TTGACGCCCCATTCACCGCAAAACAGGCGGTATGTGATTAACCACTTCACTAGTACTG 960
 DB 2711 TTGACGCCCCATTCACCGCAAAACAGGCGGTATGTGATTAACCACTTCACTAGTACTG 2770
 QY 961 TTTATTTCCGAGACAGTACTTAATCTGGCAATCTCGGCGGCGCAGTGAAGCTAAGTGG 1020
 DB 2771 TTTATTTCCGAGACAGTACTTAATCTGGCAATCTCGGCGGCGCAGTGAAGCTAAGTGG 2830
 QY 1021 ACCTTCCCGGTCAGCGGATTAACAGCGCGGCGAGTGTGATTAACCACTTCACTAGTACTG 1080
 DB 2831 ACCTTCCCGGTCAGCGGATTAACAGCGCGGCGAGTGTGATTAACCACTTCACTAGTACTG 2890
 QY 1081 CGTGGGCTTAACGATTAACAGCAGCAGTGAATTCAGTTTCGCTGCTTCCAGACTTAAAG 1140
 DB 2891 CGTGGGCTTAACGATTAACAGCAGCAGTGAATTCAGTTTCGCTGCTTCCAGACTTAAAG 2950
 QY 1141 CAGATGGCTGATTAACAGCGGCTGCTTAATACGCGCGCGGAGGAGTGAACAGC 1200
 DB 2951 CAGATGGCTGATTAACAGCGGCTGCTTAATACGCGCGCGGAGGAGTGAACAGC 3010
 QY 1201 CTGCGAGAGTGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTTGGCAGGTTTAAACCAA 1260
 DB 3011 CTGCGAGAGTGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTTGGCAGGTTTAAACCAA 3070
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAAGTTTG 1296
 DB 3071 ATCGTGAATGAAGCAGCATACCGCGCTGCAAGTTTG 3106
 RESULT 8
 AAC68297
 ID AAC68297 standard; DNA; 6116 BP.
 XX AAC68297;
 AC
 XX 20-FEB-2001 (first entry)
 DT
 XX
 XX R15/APPA plasmid coding sequence.
 DE
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig; ds.
 OS
 XX Rattus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX
 XX W0200064247-A1.
 XX
 XX 02-NOV-2000.
 PD
 XX 20-APR-2000; 2000MO-CA00430.
 PF
 XX 23-APR-1999; 99US-0130508.
 PR
 XX (UYGU-) UNIV GUELPH.
 PA

XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.

XX Sequence 17732 BP; 4719 A; 4125 C; 4168 G; 4719 T; 1 other:

Query Match 96.98; Score 1282.2; DB 21; Length 17732;

Best Local Similarly 99.38; Pred. No. 0;

Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGCTTAACCCCGCATCTGCA 60
DB 12653 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGCTTAACCCCGCATCTGCA 12712
QY 61 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGAGTGGTATGTCAGTCTCATGCT 120
DB 12713 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGAGTGGTATGTCAGTCTCATGCT 12772
QY 121 GTGCGTCTCCAAACCAAGGCGCAGCACTGATGAGATGTCACCCGAGAGCATGAGCCA 180
DB 12773 GTGCGTCTCCAAACCAAGGCGCAGCACTGATGAGATGTCACCCGAGAGCATGAGCCA 12832
QY 181 ACCTGCGCGGTAAACCTGGGTGGCTGACACCGGNGGTGGTGAAGTATTCGCTATCTC 240
DB 12833 ACCTGCGCGGTAAACCTGGGTGGCTGACACCGGNGGTGGTGAAGTATTCGCTATCTC 12892
QY 241 GGACATTACCAACCGCCAGCTGCTGAGCCAGGATTTGCTGGGAAAGAGGCTGCCG 300
DB 12893 GGACATTACCAACCGCCAGCTGCTGAGCCAGGATTTGCTGGGAAAGAGGCTGCCG 12952
QY 301 CAGTCTGCTGAGTCCGCGATTTATGCTGATGTCAGCAGGAGTACCCGTAAGAGCGCAA 360
DB 12953 CAGTCTGCTGAGTCCGCGATTTATGCTGATGTCAGCAGGAGTACCCGTAAGAGCGCAA 13012
QY 361 GCGTTCGCGCGCGCGGCTGGGACCTGATGTCATTAACCGTACTACCCAGGAGATACG 420
DB 13013 GCGTTCGCGCGCGCGGCTGGGACCTGATGTCATTAACCGTACTACCCAGGAGATACG 13072
QY 421 TCCAGTCCGATCCGTTATTTAACTCTCTAAAGCTGGCTTTGCCAAGTGAATACGCG 480
DB 13073 TCCAGTCCGATCCGTTATTTAACTCTCTAAAGCTGGCTTTGCCAAGTGAATACGCG 13132
QY 481 AACGTGACTAGCGGATCTCTAGCAGGAGGAGGATCAATGCTGACTTACCGGGCAT 540
DB 13133 AACGTGACTAGCGGATCTCTAGCAGGAGGAGGATCAATGCTGACTTACCGGGCAT 13192
QY 541 CGGCAAAAGCGGCTTTCCGGAACCTGGAACGGGTCTTAATTTCCGCAATCAACTTGTGC 600
DB 13193 CGGCAAAAGCGGCTTTCCGGAACCTGGAACGGGTCTTAATTTCCGCAATCAACTTGTGC 13252
QY 601 CTTAAAGCTGAGAAAGAGGAGGAAAGCTGTTCAATTAACGAGGATTAACCATGGAATC 660
DB 13253 CTTAAAGCTGAGAAAGAGGAGGAAAGCTGTTCAATTAACGAGGATTAACCATGGAATC 13312
QY 661 AAGGTGAGCGCGCAATGCTCTAATTAACCGGTGCGTAAAGCCCTGCGATCAATGCTGAG 720
DB 13313 AAGGTGAGCGCGCAATGCTCTAATTAACCGGTGCGTAAAGCCCTGCGATCAATGCTGAG 13372
QY 721 GAATATATTTCTCTGCAACAAGCAGGGAATGCGGAGCGGGGTGGGAAAGATACCC 780
DB 13373 GAATATATTTCTCTGCAACAAGCAGGGAATGCGGAGCGGGGTGGGAAAGATACCC 13432
QY 781 GATTACACAGTGAACACCTTCTAAGTTTGAATTAACGCGCAATTTTATTTCTTCAAA 840
DB 13433 GATTACACAGTGAACACCTTCTAAGTTTGAATTAACGCGCAATTTTATTTCTTCAAA 13492
QY 841 CGCAGCGCAGAGTGGCGCGCAGCGCGCAGCCCGTATTTGATTTGATCATGACGAGCG 900
DB 13493 CGCAGCGCAGAGTGGCGCGCAGCGCGCAGCCCGTATTTGATTTGATCATGACGAGCG 13552

QY 901 TTGACGCGCCATCCACCGCAAAAGAGCGGATGCTGACATTAACCACTTCACTG 960
DB 13553 TTGACGCGCCATCCACCGCAAAAGAGCGGATGCTGACATTAACCACTTCACTG 13612
QY 961 TTTATTCGCGGACAGATTAATCTGCAAAATCTGCGCGGCGCAGTGAAGCTCACTG 1020
DB 13613 TTTATTCGCGGACAGATTAATCTGCAAAATCTGCGCGGCGCAGTGAAGCTCACTG 13672
QY 1021 ACGCTTCGCGGTCAGCGGATTAACACCGCGCGAGGTGTAACCTGTTGAACGCTG 1080
DB 13673 ACGCTTCGCGGTCAGCGGATTAACACCGCGCGAGGTGTAACCTGTTGAACGCTG 13732
QY 1081 CGTGGGCTAAGCGATTAACAGCCAGTGAATTCAGTTTCGCTGCTTCCAGCTTACAG 1140
DB 13733 CGTGGGCTAAGCGATTAACAGCCAGTGAATTCAGTTTCGCTGCTTCCAGCTTACAG 13792
QY 1141 CAGATGCGGTGATTAAGCGCGCTGCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
DB 13793 CAGATGCGGTGATTAAGCGCGCTGCTAATTAATACGCGCGCGGAGAGGTGAACCTGACC 13852
QY 1201 CTGCGAGATGTCGAAGACCAATGCGCAGGCGCATGTTGCTTGGCAGGTTTACGCAA 1260
DB 13853 CTGCGAGATGTCGAAGACCAATGCGCAGGCGCATGTTGCTTGGCAGGTTTACGCAA 13912
QY 1261 ATCGTAATGAAGCAGCATACCGCGCGTTCG 1296
DB 13913 ATCGTAATGAAGCAGCATACCGCGCGTTCG 13948

RESULT 11

AAC68294

ID AAC68294 standard; DNA; 20623 BP.

AC 68294;

20-FEB-2001 (first entry)

XX Lamaz/APPA plasmid coding sequence.

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;

XX environmental pollution; pig; ds.

OS Mus musculus - chimeric.

OS Escherichia coli - chimeric.

XX WO200064247-A1.

XX 02-NOV-2000.

XX 20-APR-2000; 2000WO-CA00430.

XX 23-APR-1999; 99US-0130508.

XX (UYGU-) UNIV GUELPH.

PI Forsberg CW, Golovan S, Phillips JP;

XX WPI; 2000-687245/67.

XX P-PSDB; AAB36257.

XX Transgenic non-human animal for gastrointestinal tract specific

XX expression of a protein, preferably phytase, comprises a nucleic acid

XX sequence including a heterologous transgene construct encoding the

XX protein

XX Claim 56; Fig 5; 152pp; English.

XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of

transgenes containing the E. coli APPA phytase coding sequence.

CC Sequence 20623 BP; 5449 A; 4847 C; 4902 G; 5424 T; 1 other;

Query Match 96.9%; Score 1282.2; DB 21; Length 20623;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

QY 1 ATGAAGGATCTTATCCATTTTATCTCTTCGATTCGTTAACCCGCAATCTCA 60
DB 12653 ATGAAGGATCTTATCCATTTTATCTCTTCGATTCGTTAACCCGCAATCTCA 12712
QY 61 TTCGCTCAGAGTACGCGGAGCTGAGTGAAGTGTGTGATTTGTCAGTCGTCAGT 120
DB 12713 TTCGCTCAGAGTACGCGGAGCTGAGTGAAGTGTGTGATTTGTCAGTCGTCAGT 12772
QY 121 GTGGGTGCTCCACCAAGGCGCAGCACTGATGAGATGTCACCCGAGAGCATGGCCA 180
DB 12773 GTGGGTGCTCCACCAAGGCGCAGCACTGATGAGATGTCACCCGAGAGCATGGCCA 12832
QY 181 ACCTGGCCGCTAAACTGGGTTGGCTGACACCGCGNGTGGTGAAGCTAATCGCTATCTC 240
DB 12833 ACCTGGCCGCTAAACTGGGTTGGCTGACACCGCGNGTGGTGAAGCTAATCGCTATCTC 12892
QY 241 GGACATTAACCAACGCCACGCTCTGTGACCGCAGCGGATTCGTGGCGAAAAAGGCTGCCG 300
DB 12893 GGACATTAACCAACGCCACGCTCTGTGACCGCAGCGGATTCGTGGCGAAAAAGGCTGCCG 12952
QY 301 CAGTCTGGTCAAGTCCGCGATTAATGCTGATGTCGACGAGCGGTACCCGTAACCAAGCGGAA 360
DB 12953 CAGTCTGGTCAAGTCCGCGATTAATGCTGATGTCGACGAGCGGTACCCGTAACCAAGCGGAA 13012
QY 361 GCGTTCCGCGCGGCGGCTGGCAGCCTGACTGTGCAATTAACCGTACATACCCAGGAGATACG 420
DB 13013 GCGTTCCGCGCGGCGGCTGGCAGCCTGACTGTGCAATTAACCGTACATACCCAGGAGATACG 13072
QY 421 TTCGATCCCGGATCCGTTATTTAATCCTTAATAAACTGGCGTTGGCACTGATTAACCGG 480
DB 13073 TTCGATCCCGGATCCGTTATTTAATCCTTAATAAACTGGCGTTGGCACTGATTAACCGG 13132
QY 481 AACGTGACTGACGAGATCTCTCAGCAGGAGGAGGATCAATTCGTCATTAACCGGAGCAT 540
DB 13133 AACGTGACTGACGAGATCTCTCAGCAGGAGGAGGATCAATTCGTCATTAACCGGAGCAT 13192
QY 541 CGGCAAAACGGGCTTTCGGGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
DB 13193 CGGCAAAACGGGCTTTCGGGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 13252
QY 601 CTTAAAGCTGAGAAACAGAGAGAAAGCTGTTCATTAACGAGGCAATTACCTCGGAATCTC 660
DB 13253 CTTAAAGCTGAGAAACAGAGAGAAAGCTGTTCATTAACGAGGCAATTACCTCGGAATCTC 13312
QY 661 AAGGTGAGCGCGGCAATGTCTCATTAACCGGTGGGTAACTCGCATCATGTGTCAG 720
DB 13313 AAGGTGAGCGCGGCAATGTCTCATTAACCGGTGGGTAACTCGCATCATGTGTCAG 13372
QY 721 GAGATATTTTCTCTGCAACAAGCAGAGGAATGCCGAGCCGGGGTGGGAGGATCAACC 780
DB 13373 GAGATATTTTCTCTGCAACAAGCAGAGGAATGCCGAGCCGGGGTGGGAGGATCAACC 13432
QY 781 GATTCACACCAAGTGGAAACACTTGTCTAATTAACGGCGCAATTTTATTTGCTACAA 840
DB 13433 GATTCACACCAAGTGGAAACACTTGTCTAATTAACGGCGCAATTTTATTTGCTACAA 13492
QY 841 CGCAGCGCAGAGGTGGCGCGGAGCGCGCCCGTTATGATGATGATGATGATGATGATGATG 900
DB 13493 CGCAGCGCAGAGGTGGCGCGGAGCGCGCCCGTTATGATGATGATGATGATGATGATGATG 13552
QY 901 TTGAGCGCCCATTCACACGCAAAACAGCGATGATGATGATGATGATGATGATGATGATG 960
DB 13553 TTGAGCGCCCATTCACACGCAAAACAGCGATGATGATGATGATGATGATGATGATGATG 13612
QY 961 TTATATTCGCGGACAGATTAATCTGCGCAAAATCTCGGCGGCGACGTGAGACTCAACTGG 1020

```

```

DB 13613 TTATATTCGCGGACAGATTAATCTGCAAAATCTCGCGCGGCGACAGTGGACTCAACTGG 13672
QY 1021 ACGGTTCCCGGTCAGCGGATTAACACCGCGGAGGTGTAAGTGTGTTGAACGCTGG 1080
DB 13673 ACGGTTCCCGGTCAGCGGATTAACACCGCGGAGGTGTAAGTGTGTTGAACGCTGG 13732
QY 1081 CGTCGGCTAACGATTAACAGCAGTGTAGTTCGCTGCTGCTTCACACTTTACAG 1140
DB 13733 CGTCGGCTAACGATTAACAGCAGTGTAGTTCGCTGCTGCTTCACACTTTACAG 13792
QY 1141 CAGATGCTGATTAACCGCGCTGTCTAATTAATACCGCGCGGAGAGGTGAACCTGACC 1200
DB 13793 CAGATGCTGATTAACCGCGCTGTCTAATTAATACCGCGCGGAGAGGTGAACCTGACC 13852
QY 1201 CTGGCAGATGTAAGAGCAGAAATGCGCAGGAGCATGTGTGCTTGGCAGTTTACGCA 1260
DB 13853 CTGGCAGATGTAAGAGCAGAAATGCGCAGGAGCATGTGTGCTTGGCAGTTTACGCA 13912
QY 1261 ATCGTAATGAAGCAGCATACCGCGCTGACAGTTTG 1296
DB 13913 ATCGTAATGAAGCAGCATACCGCGCTGACAGTTTG 13948

```

RESULT 12

```

ID AAC68299 standard; DNA; 5421 BP.
AC AAC68299;
AC 20-FEB-2001 (first entry)
DE SV40/APPA plasmid coding sequence.
DE Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
DE environmental pollution; pig; ds.
KW Rhesus macaque polyoma virus - chimeric.
OS Escherichia coli - chimeric.
PN WO200064247-A1.
PD 02-NOV-2000.
XX 20-APR-2000; 2000WO-CA00430.
XX 23-APR-1999; 99US-0130508.
XX (UIGU-) UNIV GUELPH.
PA Forsberg CW, Golovan S, Phillips JP;
PI MPI: 2000-687245/67.
DR P-PSDB; AAB36262.
XX Transgenic non-human animal for gastrointestinal tract specific
XX expression of a protein, preferably phytase, comprises a nucleic acid
XX sequence including a heterologous transgene construct encoding the
XX protein.
PS Claim 56; Fig 22; 152pp; English.
XX The present invention provides transgenic animals which produce desired
XX proteins, in this case pigs which expresses phytase in the salivary
XX gland. Low phytase production levels result in phytate in the diet being
XX excreted and causing phosphorus contamination in water, as well as
XX reducing the growth of animals. The invention provides a number of
XX transgenes containing the E. coli APPA phytase coding sequence.
SQ Sequence 5421 BP; 1413 A; 1321 C; 1331 G; 1355 T; 1 other;
Query Match 96.0%; Score 1270.2; DB 21; Length 5421;
Best Local Similarity 99.3%; Pred. No. 0;

```


CC activity. The phosphatase fragment is created by treating the phosphatase
CC with a protease. The invention includes an Escherichia coli nucleotide
CC sequence which encodes an acidic phosphatase appa2 which has improved
CC phosphatase activity. The appa gene of Escherichia coli was originally
CC defined as E. coli periplasmic phosphohydrolase phosphohydrolase,
CC although it was not previously known to have phytase activity. The
CC phosphatase fragment can be used in animal feed, to improve the
CC accessibility of phosphate to the animal. The phosphatase gene which
CC encodes the protein can be used in the production of phytase. The present
CC sequence represents the appa2 coding sequence.

XX Sequence 1489 BP; 362 A; 384 C; 403 G; 340 T; 0 other:

Query Match 95.5%; Score 1263.8; DB 21; Length 1489;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1280; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

QY 1 ATGAAACGATCTTAAATCCCATTTTATCTTCTGATCCGTTAACCCGCAATGCA 60
DB 182 ATGAAACGATCTTAAATCCCATTTTATCTTCTGATCCGTTAACCCGCAATGCA 241
QY 61 TTGCTCAGAGTAGAGCCGAGCTAGAGTGAAGTGTGATTCAGTGCATAGGT 120
DB 242 TTGCTCAGAGTAGAGCCGAGCTAGAGTGAAGTGTGATTCAGTGCATAGGT 301
QY 121 GTGCGTCTCCAAACCAAGGCCACCACTGATGAGATGTACACCCAGACGATGGCA 180
DB 302 GTGCGTCTCCAAACCAAGGCCACCACTGATGAGATGTACACCCAGACGATGGCA 361
QY 181 ACCGTGCGGTGAAACCTGGGTGCTGACACCGGNGGTGTGAGCTATATCTC 240
DB 362 ACCGTGCGGTGAAACCTGGGTGCTGACACCGGNGGTGTGAGCTATATCTC 421
QY 241 GGACATTAACCAACGCCAGCTGTGTAGCAGCAGGATTTGGGCAAAAAGGGTGGCCG 300
DB 422 GGACATTAACCAACGCCAGCTGTGTAGCAGCAGGATTTGGGCAAAAAGGGTGGCCG 481
QY 301 CAGTGTGTCAGTGTGCGGATTTATGCTGATGTGACAGACGCTAACCCGTAACAGGCGAA 360
DB 482 CAGTGTGTCAGTGTGCGGATTTATGCTGATGTGACAGACGCTAACCCGTAACAGGCGAA 541
QY 361 GCCCTGCGCGCGCGGCGGACCTGACTGTGCAATACCCGTAACAGGCGAA 420
DB 542 GCCCTGCGCGCGCGGCGGACCTGACTGTGCAATACCCGTAACAGGCGAA 601
QY 421 TCCAGTCCGATCCGTTATTTAATCTCTAATAAACTGGCTTCCCACTGGATACGCG 480
DB 602 TCCAGTCCGATCCGTTATTTAATCTCTAATAAACTGGCTTCCCACTGGATACGCG 661
QY 481 AACGTGACTGACGGATCTCTACAGGAGGAGGATCAATTTGCTGACTTTACCGGCAT 540
DB 662 AACGTGACTGACGGATCTCTACAGGAGGAGGATCAATTTGCTGACTTTACCGGCAT 721
QY 541 CGGCAAAACGGCGTTCCGCAACTGGAACGGGTGTTAATTTCCGCAATCAACTTGTGC 600
DB 722 CGGCAAAACGGCGTTCCGCAACTGGAACGGGTGTTAATTTCCCAATTAACCTTGTGC 781
QY 601 CTTAAAGGTGAGAAACAGAGAAAGCTGTTCAATTAACGAGGATTAACCATGGAATC 660
DB 782 CTTAAAGGTGAGAAACAGAGAAAGCTGTTCAATTAACGAGGATTAACCATGGAATC 841
QY 661 AAGGTGAGCCCGCAATGTTCTCAATTAACGGGTGGAAGCTTCGATCAATGCTGACG 720
DB 842 AAGGTGAGCCCGCAATGTTCTCAATTAACGGGTGGAAGCTTCGATCAATGCTGACG 901
QY 721 GAGATATTTCTCTCGCAACAAGCACAGGAAATGCGGAGCCGGGTGGGGAAGATCAAC 780
DB 902 GAGATATTTCTCTCGCAACAAGCACAGGAAATGCGGAGCCGGGTGGGGAAGATCACT 961
QY 781 GATTACACACAGTGAACACCTTGTGAAGTTTGATTAACGGGCAATTTTATTTACTACAA 840
DB 962 GATTACACACAGTGAACACCTTGTGAAGTTTGATTAACGGGCAATTTTATTTACTACAA 1021

```

```

QY 841 CGACGCCAGAGGTGTCGCCGACGCCGCCACCCGTTATTTGATTTGATGATGCAACG 900
DB 1022 CGCAGCCAGAGGTGTCGCCGACGCCGCCACCCGTTATTTGATTTGATGATGCAACG 1081
QY 901 TTGACGCCCGCATTCACCGCAAAACAGCGGATGTGAGATTAACCACTTCAGTACTG 960
DB 1082 TTGACGCCCGCATTCACCGCAAAACAGCGGATGTGAGATTAACCACTTCAGTACTG 1141
QY 961 TTTATTCGCCGACAGTAACTAATTCGCAATTCGCGCGCGCATGAGACTGCACTG 1020
DB 1142 TTTATTCGCCGACAGTAACTAATTCGCAATTCGCGCGCGCATGAGACTGCACTG 1201
QY 1021 ACCTTCCCGGTACAGCGGATTAACACGCCCGCAGGTGTGATGCTGTTGAACGCTG 1080
DB 1202 ACCTTCCCGGTACAGCGGATTAACACGCCCGCAGGTGTGATGCTGTTGAACGCTG 1261
QY 1081 CGTGGCTTAACGCAACAGCAGTGTGATTCAGTTCGCGGTGCTTCCAGCTTACAG 1140
DB 1262 CGTGGCTTAACGCAACAGCAGTGTGATTCAGTTCGCGGTGCTTCCAGCTTACAG 1321
QY 1141 CAGATGCGGTAAACAGCCGCTGTCAATTAATACGCCCGCGGAGAGTGAACCTGACC 1200
DB 1322 CAGATGCGGTAAACAGCCGCTGTCAATTAATACGCCCGCGGAGAGTGAACCTGACC 1381
QY 1201 CTGGCAGAGTGAAGAGCGAATGCGCAGGCGATGTGTTGCGCAGGTTTAAACGCA 1260
DB 1382 CTGGCAGAGTGAAGAGCGAATGCGCAGGCGATGTGTTGCGCAGGTTTAAACGCA 1441
QY 1261 ATCGTGAATGAAGCAGCATCCGCGGTGCTTGAATGCTCAATC 1308
DB 1442 ATCGTGAATGAAGCAGCATCCGCGGTGCTTGAATGCTCAATC 1489

```

RESULT 14

AAD06831
ID AAD06831 standard; DNA; 1486 BP.

XX AAD06831;

DT 06-AUG-2001 (first entry)

DE E. coli acid phosphatase/phytase (appa) DNA.

KW Acid phosphatase-phytase; appa; enzymatic activity; food additive;

KV animal feed; monoester phosphate; cereal food; human food; ds.

XX Escherichia coli.

XX OS Escherichia coli.

FH Key. Location/Qualifiers

FT CDS 16..108

FT /tag- a "Peptide encoded by shorter ORF of appa DNA"

FT /product- "E. coli appa protein"

FT /transl_except- (pos:512..514, aa:Ala)

FT /transl_except- (pos:629..631, aa:Asn)

FT complement (245..267)

FT /tag- c "Forward PCR primer E2"

FT /bound_molety- "Forward PCR primer E2"

FT /tag- d "Reverse PCR primer A1"

FT /bound_molety- "Reverse PCR primer A1"

FT /tag- e "Forward PCR primer P2"

FT /bound_molety- "Forward PCR primer P2"

FT /tag- f "Reverse PCR primer P3"

FT /bound_molety- "Reverse PCR primer P3"

FT /tag- g "Reverse PCR primer K2"

FT /bound_molety- "Reverse PCR primer K2"

PN W0200136607-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 17-NOV-2000; 2000MO-US31622.
 XX
 PR 18-NOV-1999; 99US-0166179.
 XX
 PA (CORR) CORNELL RES FOUND INC.
 XX
 PI Lei X;
 XX
 DR WPI: 2001-367572/38.
 DR P-PSDB: AAE02631, AAE02633.
 XX
 PT Mutated acid phosphatase/phytase from *Escherichia coli* has improved
 PT enzymatic activity compared to the wild type and is useful as a food
 PT additive, particularly for animal feeds
 XX
 PS Claim 10; Page 7-9; 56pp; English.
 XX
 CC The present sequence is a DNA encoding wild type acid phosphatase/phytase
 CC (appa) from *Escherichia coli*. The wild type appa is mutated to enhance
 CC its enzymatic activity. Phytases, a specific group of monoester from
 CC phosphates, are required to initiate the release of phosphate from
 CC phytate, the major storage of phosphate in cereal foods or feeds. The
 CC mutant acid phosphatase/phytase is added to animal feed as a food
 CC additive to improve uptake of phosphate and zinc from the diet. The
 CC phytase can also be added to human food. The mutant phytase is more heat
 CC stable than phytase isolated from *Aspergillus niger* and safer for use
 CC in human food manufacture.
 CC
 SO Sequence 1486 BP; 367 A; 380 C; 403 G; 336 T; 0 other;

Query Match 95.5%; Score 1263; DB 22; Length 1486;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1275; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

```

QY 1 ATGAAGGAGATCTTAATCCATTCTTCTGATCCGTTAAACCCCGAATCTGCA 60
DB 188 ATGAAGGAGATCTTAATCCATTCTTCTGATCCGTTAAACCCCGAATCTGCA 247
QY 61 TTCGCTCAGAGTACCCGAGCTGGAAGTGAAGTGTGATGTCAGTGTCTGCT 120
DB 248 TTCGCTCAGAGTACCCGAGCTGGAAGTGAAGTGTGATGTCAGTGTCTGCT 307
QY 121 GTGCGTGTCTCAACCAAGGCCACGCACTGATGCAAGATGTCAACCCAGAGCAGTGGCA 180
DB 308 GTGCGTGTCTCAACCAAGGCCACGCACTGATGCAAGATGTCAACCCAGAGCAGTGGCA 367
QY 181 ACCTGGCCGCTAAACTGGGTTGGTGAACACCGCGGTGGTGAAGTATCCGCTATCTC 240
DB 368 ACCTGGCCGCTAAACTGGGTTGGTGAACACCGCGGTGGTGAAGTATCCGCTATCTC 427
QY 241 GGACATTAACCAAGCCAGCGTCTGTAGCGACGATGCTGGCGAAAAAGGCTGCCG 300
DB 428 GGACATTAACCAAGCCAGCGTCTGTAGCGACGATGCTGGCGAAAAAGGCTGCCG 487
QY 301 CAGTGTGCTCAGTGGCGATTAATGCTGATGTGCAAGAGCGTACCCGTTAAACAGGCGAA 360
DB 488 CAGTGTGCTCAGTGGCGATTAATGCTGATGTGCAAGAGCGTACCCGTTAAACAGGCGAA 547
QY 361 GCGTTCGCGCGCGGGGTGGCAGCTGACATGTCGAATTAACGTTACCCAGCAGATAG 420
DB 548 GCGTTCGCGCGCGGGGTGGCAGCTGACATGTCGAATTAACGTTACCCAGCAGATAG 607
QY 421 TCCAGTCCCGATCCGTTATTTATCTCTAAACAACTGGCGTTGCCAATGTGATAAGCG 480
DB 608 TCCAGTCCCGATCCGTTATTTATCTCTCTAAACAACTGGCGTTGCCAATGTGATAAGCG 667
QY 481 AACGTGATGACGCGATCTCTACAGAGGCGAGAGGCTCAATGCTGACTTTACCGGGCAT 540
DB 668 AACGTGATGACGCGATCTCTCTACAGAGGCGAGAGGCTCAATGCTGACTTTACCGGGCAT 727
  
```

```

QY 541 CGGCAACAGCGCGTTTCGGCACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGCG 600
DB 728 CGGCAACAGCGCGTTTCGGCACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGCG 787
QY 601 CTTAAACGTGAGAAACAGCAGCAAGAGCTGTGATTAACGACAGCATTTACCATCGAATCTC 660
DB 788 CTTAAACGTGAGAAACAGCAGCAAGAGCTGTGATTAACGACAGCATTTACCATCGAATCTC 847
QY 661 AAGGTGAGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCCTGCAATCAATGCTGACG 720
DB 848 AAGGTGAGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCCTGCAATCAATGCTGACG 907
QY 721 GAGATTTCTCTCCGCAACAGACAGCGGAATGCGGACCGGGGGGGGAAGGATACAC 780
DB 908 GAAATATTTCTCTCCGCAACAGACAGCGGAATGCGGACCGGGGGGGGAAGGATACAC 967
QY 781 GATTCACACAGTGAACACCTGCTAAGTTCATTAACGCGCAATTTATTTGCTACAA 840
DB 968 GATTCACACAGTGAACACCTGCTAAGTTCATTAACGCGCAATTTATTTGCTACAA 1027
QY 841 CGCAGCGCAGAGGTTCCCGCAGCCGCGCACCCGTTATTTGATTTGATGACGACG 900
DB 1028 CGCAGCGCAGAGGTTCCCGCAGCTCCGCGCACCCGTTATTTGATTTGATGACGACG 1087
QY 901 TTGAGCGCGCATCCACCGCAAAACAGGGGTATGTTGATTAACCACTTCAGTACG 960
DB 1088 TTGAGCGCGCATCCACCGCAAAACAGGGGTATGTTGATTAACCACTTCAGTACG 1147
QY 961 TTTATTTCCGCGACAGCATTAATCTGCAAAATCTCGCGGCGCACTGGAGGCTCACTG 1020
DB 1148 TTTATTTCCGCGACAGCATTAATCTGCAAAATCTCGCGGCGCACTGGAGGCTCACTG 1207
QY 1021 AGCGTTCCCGGTACCGCGATTAACAGCCGCGCAGGTGTGTAAGTGTGTTGAACGCTG 1080
DB 1208 AGCGTTCCCGGTACCGCGATTAACAGCCGCGCAGGTGTGTAAGTGTGTTGAACGCTG 1267
QY 1081 CGTCCGCTAAGCGATTAACAGCCGCGATTAACAGGTTGCTGCTGCTTCCAGACTTACAG 1140
DB 1268 CGTCCGCTAAGCGATTAACAGCCGCGATTAACAGGTTGCTGCTGCTTCCAGACTTACAG 1327
QY 1141 CAGATCGGTATTAACAGCCGCGTGTATTAATATACCCCGGAGAGGTGAATGAC 1200
DB 1328 CAGATCGGTATTAACAGCCGCGTGTATTAATATACCCCGGAGAGGTGAATGAC 1387
QY 1201 CTGGCAGATGTGAAGCGGAATGCGCAGGCGATGTGCTGTCGACAGTTTACGCA 1260
DB 1388 CTGGCAGATGTGAAGCGGAATGCGCAGGCGATGTGCTGTCGACAGTTTACGCA 1447
QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTG 1296
DB 1448 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTG 1483
  
```

RESULT 15
 AAD06832 standard; DNA; 1486 BP.
 AC AAD06832;
 DT 06-AUG-2001 (first entry)
 XX
 DE E. coli acid phosphatase/phytase (appa) mutant DNA.
 KW Acid phosphatase-phytase; appa; enzymatic activity; food additive;
 OS animal feed; monoester phosphate; cereal food; human food; mutant; ds.
 OS *Escherichia coli*.
 XX Synthetic.
 FH Key Location/Qualifiers
 FT CDS 16..108
 FT /*tag= a

Search completed: June 12, 2003, 10:33:46
Job time : 373 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 08:17:53 ; Search time 235 Seconds
(without alignments)
7879.096 Million cell updates/sec

Title: US-09-777-566A-1
Perfect score: 1323
Sequence: 1 atgaagcagctatccttacc.....atcacatcacatcactaa 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCr_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCrUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Result					ID	Description
No.	Score	Match	Length	DB		
1	1322	99.9	1323	9	US-10-034-985-1	Sequence 1, Appl1
2	1322	99.9	1323	10	US-09-777-566A-1	Sequence 1, Appl1
3	1322	99.9	1323	10	US-09-866-379-1	Sequence 1, Appl1
4	1287.8	97.3	1901	10	US-09-866-379-7	Sequence 7, Appl1
5	1286.2	97.2	1901	10	US-09-866-379-5	Sequence 5, Appl1
6	1286.2	97.2	1901	10	US-09-866-379-9	Sequence 9, Appl1
7	1283	97.0	1901	10	US-09-866-379-6	Sequence 6, Appl1
8	1263.8	95.5	1489	9	US-10-266-041-9	Sequence 9, Appl1
9	218.8	16.5	1325	9	US-10-021-723A-3	Sequence 3, Appl1
10	215.6	16.3	1325	9	US-10-021-723A-1	Sequence 1, Appl1
11	206.6	15.6	1326	9	US-10-021-723A-11	Sequence 11, Appl1
12	159.6	12.1	1431	9	US-10-021-723A-9	Sequence 9, Appl1
13	73	5.5	1230	9	US-10-021-723A-7	Sequence 7, Appl1
14	64.8	4.9	1266	9	US-10-021-723A-5	Sequence 5, Appl1
15	36.4	2.8	802	9	US-10-184-644-312	Sequence 312, App
16	36.4	2.8	802	9	US-10-184-634-312	Sequence 312, App
17	35.4	2.7	2781	9	US-10-092-154-1364	Sequence 1364, App
18	35.4	2.7	2781	10	US-09-764-847-1364	Sequence 1364, App
19	2.6	867	9	US-10-123-155-20	Sequence 20, Appl1	

c	20	33.8	2.6	54	9	US-10-021-723A-47	Sequence 47, Appl
	21	33.8	2.6	450	9	US-09-796-692-7902	Sequence 7902, Ap
	22	33.8	2.6	450	9	US-10-040-862-7902	Sequence 7902, Ap
	23	33.8	2.6	508	9	US-09-796-692-8272	Sequence 8272, Ap
	24	33.8	2.6	508	9	US-10-040-862-8272	Sequence 8272, Ap
	25	33.4	2.5	974	10	US-09-770-445-281	Sequence 281, Appl
	26	33	2.5	790	9	US-10-123-155-204	Sequence 204, Appl
	27	32.6	2.5	55	9	US-10-021-723A-22	Sequence 22, Appl
	28	32.6	2.5	14800	10	US-09-954-456-1601	Sequence 1601, Ap
	29	32.4	2.4	464	9	US-09-918-995-12838	Sequence 12838, A
	30	32.4	2.4	607	9	US-10-123-155-344	Sequence 344, Appl
	c	31	32.4	888	9	US-10-123-155-544	Sequence 544, Appl
		32	32.2	162	9	US-09-828-523A-97	Sequence 97, Appl
		33	32.2	162	9	US-09-966-521-89	Sequence 89, Appl
		34	32.2	401	10	US-09-969-708-272	Sequence 272, Appl
		35	32.2	401	10	US-09-880-107-1305	Sequence 1305, Ap
		36	32.2	457	9	US-10-123-155-236	Sequence 236, Appl
		37	31.8	614	10	US-09-974-300-2058	Sequence 2058, Ap
	c	38	31.8	1049	9	US-10-123-155-358	Sequence 358, Appl
	c	39	31.8	7558	9	US-10-091-458-52	Sequence 52, Appl
	c	40	31.6	3209	10	US-09-925-301-474	Sequence 474, Appl
	c	41	31.6	8460	9	US-10-237-271-2	Sequence 2, Appl1
	c	42	31.4	625	9	US-10-123-155-70	Sequence 70, Appl
	c	43	31.4	2890	10	US-09-881-752A-361	Sequence 361, Appl
	44	31.2	2.4	1509	9	US-09-828-523A-89	Sequence 89, Appl
	45	31.2	2.4	1509	9	US-09-966-521-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-10-034-985-1
Sequence 1, Application US/10034985
Publication No. US20030049815A1
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHRYTASE
FILE REFERENCE: 09010/029003
CURRENT APPLICATION NUMBER: US/10/034,985
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: US/09/580,515
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 09/291,931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 08/910,798
PRIOR FILING DATE: 1997-08-13
PRIOR APPLICATION NUMBER: 09/259,214
PRIOR FILING DATE: 1999-03-01
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..
NAME/KEY: misc_feature
LOCATION: (1)..
OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1
Query Match 99.9%; Score 1322; DB 9; Length 1323;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGAAGCAGATCTTATCCATTTTATCTCTGATTCGTTAACCCGCAATCGCA 60
DB 1 ATGAAGCAGATCTTATCCATTTTATCTCTGATTCGTTAACCCGCAATCGCA 60
OY 61 TTCGCTGAGAGTGGAGCGGAGCTGAAGTGGATTCAGTCGTCAGTGT 120

|||||
Db 301 CAGTCGTCAGTCGCGATTAATGCTGATGTCGAGAGAGCTACCCGTAATAACAGCGAA 360
QY 361 GCCTTGCGCGCGCGGCTGGCAGCTGACTGCAATAACCGTACATACCGAGCATAGC 420
Db 361 GCCTTGCGCGCGCGGCTGGCAGCTGACTGCAATAACCGTACATACCGAGCATAGC 420
QY 421 TCAGTCGCGATCCGTTATTAATCCGTAATAACCGGCTTGCCAACTGAGTAACGCG 480
Db 421 TCAGTCGCGATCCGTTATTAATCCGTAATAACCGGCTTGCCAACTGAGTAACGCG 480
QY 481 AACGTCAGTCAGCGATGCTGCTGACAGAGGAGGCTCAATGCTGACTTACCGGCAAT 540
Db 481 AACGTCAGTCAGCGATGCTGCTGACAGAGGAGGCTCAATGCTGACTTACCGGCAAT 540
QY 541 CGGCAAAAGCGCGTTTGCGCAACTGGAAGGCGGCTCTTAATTTCCGCAATCAACTGTGC 600
Db 541 CGGCAAAAGCGCGTTTGCGCAACTGGAAGGCGGCTCTTAATTTCCGCAATCAACTGTGC 600
QY 601 CTTAAAGCTGAGAAACAGAGCAAGCACTGCTCAATTAAGCGAGCATTAACATGGAATC 660
Db 601 CTTAAAGCTGAGAAACAGAGCAAGCACTGCTCAATTAAGCGAGCATTAACATGGAATC 660
QY 661 AAGGTAGCGCGCAGCAATGCTGCTCAATTAAGCGAGCATTAACATGGAATC 720
Db 661 AAGGTAGCGCGCAGCAATGCTGCTCAATTAAGCGAGCATTAACATGGAATC 720
QY 721 GAGATTTTCTCTGTCGCAACAGCAGGAGGATCCGCGAGCGGCGGAGGAGATCACC 780
Db 721 GAGATTTTCTCTGTCGCAACAGCAGGAGGATCCGCGAGCGGCGGAGGAGATCACC 780
QY 781 GATTACACAGAGGAAACCTGCTCAATTAAGCGAGCATTAACATGGAATC 840
Db 781 GATTACACAGAGGAAACCTGCTCAATTAAGCGAGCATTAACATGGAATC 840
QY 841 CGCAGCGCAGAGGTTGCGCGCAGCGCGCGCAGCGCGGCTTAATTTGATTTGATGCGCAGC 900
Db 841 CGCAGCGCAGAGGTTGCGCGCAGCGCGCGCAGCGCGGCTTAATTTGATTTGATGCGCAGC 900
QY 901 TTGACGCGCCATTCACCGCAAAACAGGCGTATGCTGATTAACCTTACCTGATGACTG 960
Db 901 TTGACGCGCCATTCACCGCAAAACAGGCGTATGCTGATTAACCTTACCTGATGACTG 960
QY 961 TTATTTGCGGACAGCATTAATCTGCAAAATCTGCGGCGGCGCAGCTGAGCTCACTG 1020
Db 961 TTATTTGCGGACAGCATTAATCTGCAAAATCTGCGGCGGCGCAGCTGAGCTCACTG 1020
QY 1021 ACGCTTCCGCGTCAGCGGATTAACAGCGCGCAGGCTGATGACTGCTGTTGAACGCTG 1080
Db 1021 ACGCTTCCGCGTCAGCGGATTAACAGCGCGCAGGCTGATGACTGCTGTTGAACGCTG 1080
QY 1081 CGTCGCGCTAAGCATTAACAGCGCGGATTAACAGCTTCCGCTTCCAGACTTACAG 1140
Db 1081 CGTCGCGCTAAGCATTAACAGCGCGGATTAACAGCTTCCGCTTCCAGACTTACAG 1140
QY 1141 CAGATCGTGATTAACAGCGCGGCTGCTCAATTAACCGCGCGGAGAGGAGTAACGAGC 1200
Db 1141 CAGATCGTGATTAACAGCGCGGCTGCTCAATTAACCGCGCGGAGAGGAGTAACGAGC 1200
QY 1201 CTGCGAGGATGTAAGAGGAAATGCGCAGGCGCATGTTGCTTGGCAGGTTTACGAA 1260
Db 1201 CTGCGAGGATGTAAGAGGAAATGCGCAGGCGCATGTTGCTTGGCAGGTTTACGAA 1260
QY 1261 ATCGTAAATGAACAGCAGCATACCGGCGTGTGAGTTGATGATCATCACCATCAATCAC 1320
Db 1261 ATCGTAAATGAACAGCAGCATACCGGCGTGTGAGTTGATGATCATCACCATCAATCAC 1320
QY 1321 TAA 1323
Db 1321 TAA 1323

US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ. ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1
Query Match 99.9%; Score 1322; DB 10; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGCCGTAACCCCGCATCTGCA 60
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGCCGTAACCCCGCATCTGCA 60
QY 61 TTGCTGAGAGTAGCGGAGCTGAAGCTGGAAGTGTGATGTCAGTCTCATGATGT 120
Db 61 TTGCTGAGAGTAGCGGAGCTGAAGCTGGAAGTGTGATGTCAGTCTCATGATGT 120
QY 121 GTGCGTCTCAACCAAGGCGCAGCACTGATGAGGATGTCACCCGAGCATGGCCA 180
Db 121 GTGCGTCTCAACCAAGGCGCAGCACTGATGAGGATGTCACCCGAGCATGGCCA 180
QY 181 ACCTGCGGCTAAACAGCTGCTGCTGACACCGCGNGTGTGAGCTAATCGGCTATCTC 240
Db 181 ACCTGCGGCTAAACAGCTGCTGCTGACACCGCGNGTGTGAGCTAATCGGCTATCTC 240
QY 241 GGAACATTAACCAAGCGCGCTGCTGCTGAGCGAGGATTCGTCGCAAAAAAGGCGTCCG 300
Db 241 GGAACATTAACCAAGCGCGCTGCTGCTGAGCGAGGATTCGTCGCAAAAAAGGCGTCCG 300
QY 301 CAGTCTGCTAGGTCGCGATTAATGCTGATGTCGAGAGGCTAACCCGTAATAACAGCGGAA 360
Db 301 CAGTCTGCTAGGTCGCGATTAATGCTGATGTCGAGAGGCTAACCCGTAATAACAGCGGAA 360
QY 361 GCCTTGCGCGCGCGGCTGGCAGCTGACTGCTGCAATAACCGTACATACCGAGGAGATAGC 420
Db 361 GCCTTGCGCGCGCGGCTGGCAGCTGACTGCTGCAATAACCGTACATACCGAGGAGATAGC 420
QY 421 TCAGTCGCGATCCGTTATTAATCCGTAATAACCGGCTTGCCAACTGAGTAACGCG 480

```

Db      421  TCCAGTCCCATCCGTTATTTAAATCCCTAAAAAAGCGCTTGGCCAAAGCATGAACGCG 480
Oy      481  AACGTGACGAGGATCCTCAGACGGGAGGAGGATGATGCTGATTAACGGGCAAT 540
Db      481  AACGTGACGAGGATCCTCAGACGGGAGGAGGATGATGCTGATTAACGGGCAAT 540
Oy      541  CGGCAAAAGCGGCTTTCGCAAGTGAACGGGCTTAATTTTCGCAATCAAACTTGTG 600
Db      541  CGGCAAAAGCGGCTTTCGCAAGTGAACGGGCTTAATTTTCGCAATCAAACTTGTG 600
Oy      601  CTTAAACGTGAGAAACAGGAGGAAAGCTGTTCAATTAACGAGGATTAACATGGAATC 660
Db      601  CTTAAACGTGAGAAACAGGAGGAAAGCTGTTCAATTAACGAGGATTAACATGGAATC 660
Oy      661  AAGGTGAGCGCCCAATATGCTCATTTAACGGGTGAGGAGGCTGATCAATCAATGCTACG 720
Db      661  AAGGTGAGCGCCCAATATGCTCATTTAACGGGTGAGGAGGCTGATCAATCAATGCTACG 720
Oy      721  GAGATATTTCTCTGCAACAGCAGAGGAATGCCGAGCGGGGTGGGAAGGATCAC 780
Db      721  GAGATATTTCTCTGCAACAGCAGAGGAATGCCGAGCGGGGTGGGAAGGATCAC 780
Oy      781  GATTACACCAAGTGAACACACTTGTGTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 840
Db      781  GATTACACCAAGTGAACACACTTGTGTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 840
Oy      841  CGCAGCGCAGAGGTGGCCGCGAGCGCGGCGCCGCTTATTTGATTTGATCATGCGACGG 900
Db      841  CGCAGCGCAGAGGTGGCCGCGAGCGCGGCGCCGCTTATTTGATTTGATCATGCGACGG 900
Oy      901  TTGACGCCCATCCACCGCAAAAACAGCGCTATGCTGATGACATTAACCCACTTCAGTACTG 960
Db      901  TTGACGCCCATCCACCGCAAAAACAGCGCTATGCTGATGACATTAACCCACTTCAGTACTG 960
Oy      961  TTTATTTGCGGAGACGATTAATCTGCGCAAAATCTGCGCGCGGCGATGAGGAGCTCAATG 1020
Db      961  TTTATTTGCGGAGACGATTAATCTGCGCAAAATCTGCGCGCGGCGATGAGGAGCTCAATG 1020
Oy      1021  ACGCTTCCCGGTGACGCGGATTAACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db      1021  ACGCTTCCCGGTGACGCGGATTAACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Oy      1081  CGTCCGCTTAAGCGATTAACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db      1081  CGTCCGCTTAAGCGATTAACACGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Oy      1141  CAGATGCGGTATTAACAGCGCGGCTGTCTTAATTAATGCGCGCGCGGAGGAGGAGGAGGAG 1200
Db      1141  CAGATGCGGTATTAACAGCGCGGCTGTCTTAATTAATGCGCGCGCGGAGGAGGAGGAGGAG 1200
Oy      1201  CTGGCAGAGATGTGAAGAGCAAAATGCGCAGGCGATGTTCTGTTGCGAGGTTTTCACGAA 1260
Db      1201  CTGGCAGAGATGTGAAGAGCAAAATGCGCAGGCGATGTTCTGTTGCGAGGTTTTCACGAA 1260
Oy      1261  ATGCTGAATTAACAGCAGGATTAACGCGCGGCTGAGATTTGAGATCTATCAACATCAAC 1320
Db      1261  ATGCTGAATTAACAGCAGGATTAACGCGCGGCTGAGATTTGAGATCTATCAACATCAAC 1320
Oy      1321  TAA 1323
Db      1321  TAA 1323

```

RESULT 4

```

; US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin

```

```

; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match      97.3%; Score 1287.8; DB 10; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy      1  ATGAACGATCTTAATCCCATTTTATCTCTTCTGATTCGTTAAACCGGCAATCTGCA 60
Db      188  ATGAACGATCTTAATCCCATTTTATCTCTTCTGATTCGTTAAACCGGCAATCTGCA 247
Oy      61  TTGCTCAGAGTGAGCGGAGCTGAGCTGAGGAAAGTGTGATTTGATGCTGATGCT 120
Db      248  TTGCTCAGAGTGAGCGGAGCTGAGGAAAGTGTGATTTGATGCTGATGCTGATGCT 307
Oy      121  GTGCGTCTCAACCAAGGCGCAGCAGCACTGATGAGGATGATCAACCCGAGCGATGGCA 180
Db      308  GTGCGTCTCAACCAAGGCGCAGCAGCACTGATGAGGATGATCAACCCGAGCGATGGCA 367
Oy      308  GTGCGTCTCAACCAAGGCGCAGCAGCACTGATGAGGATGATCAACCCGAGCGATGGCA 367
Db      181  ACCTGCGCGGTAAACTGGGTTGGCTGACACCGCGGNGTGGTGAATCGCCTATCTC 240
Oy      368  ACCTGCGCGGTAAACTGGGTTGGCTGACACCGCGGNGTGGTGAATCGCCTATCTC 427
Db      241  GGACATTACCAAGCGCGGCTGTGTAGCCGAGGATTTGCTGGGGAAGAGGCTGGCCG 300
Oy      428  GGACATTACCAAGCGCGGCTGTGTAGCCGAGGATTTGCTGGGGAAGAGGCTGGCCG 487
Db      301  CAGTCTGTGAGGCGGAGTATTAATGCTGATGTGACGAGGATACCCGTTAAACAGGCGAA 360
Oy      488  CAGTCTGTGAGGCGGAGTATTAATGCTGATGTGACGAGGATACCCGTTAAACAGGCGAA 547
Db      361  GCGTTCCGCGCGGCGTGGACCTGACTGTGCAATTAACGCTACATACCGAGGAGATAG 420
Oy      548  GCGTTCCGCGCGGCGTGGACCTGACTGTGCAATTAACGCTACATACCGAGGAGATAG 607
Db      421  TCCAGTCCGATCCGTTATTTAATCCTCTTAATAAACTGGCGTTTCCCAATGGAATAGCG 480
Oy      608  TCCAGTCCGATCCGTTATTTAATCCTCTTAATAAACTGGCGTTTCCCAATGGAATAGCG 667
Db      481  AACGTGACGAGGATCCTCAGACGGGAGGAGGAGGATGATGCTGATTAACGGGCAAT 540
Oy      668  AACGTGACGAGGATCCTCAGACGGGAGGAGGAGGATGATGCTGATTAACGGGCAAT 727
Db      541  CGGCAAAAGCGGCTTTCGCAAGTGAACGGGCTTAATTTTCGCAATCAAACTTGTG 600
Oy      728  CGGCAAAAGCGGCTTTCGCAAGTGAACGGGCTTAATTTTCGCAATCAAACTTGTG 787

```


Qy 841 CGACGCCAGAGTTGGCCCGCAGCCGCCACCCGTTATTTGATTTGATCAGCGAGCG 900
1028 CGACGCCAGAGTTGGCCCGCAGCCGCCACCCGTTATTTGATTTGATCAGCGAGCG 1087
Qy 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGATGACATTACCACTTCAGTCTG 960
1088 TTGACGCCCATTCACCGCAAAAACAGCGGTATGATGACATTACCACTTCAGTCTG 1147
Qy 961 TTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1020
1148 TTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1207
Qy 1021 ACCTTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1080
1208 ACCTTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1267
Qy 1081 CGTGGGTAAAGGATTAACAGCCAGCGGCACTGAGTTCAGTTCGCTTCAGAGCTTACAG 1140
1268 CGTGGGTAAAGGATTAACAGCCAGCGGCACTGAGTTCAGTTCGCTTCAGAGCTTACAG 1327
Qy 1141 CAGATGCGGTAAAGGATTAACAGCCAGCGGCACTGAGTTCAGTTCGCTTCAGAGCTTACAG 1200
1328 CAGATGCGGTAAAGGATTAACAGCCAGCGGCACTGAGTTCAGTTCGCTTCAGAGCTTACAG 1387
Qy 1201 CTGGCAGGATGTAAGAGCGAAATGCGCAGGCGCATGTTGCTGTCGTCGAGCTTTTACGCAA 1260
1388 CTGGCAGGATGTAAGAGCGAAATGCGCAGGCGCATGTTGCTGTCGTCGAGCTTTTACGCAA 1447
Qy 1261 ATGCTGAATGAAGCAGCATACCGCGCTGAGTTCGAT 1300
1448 ATGCTGAATGAAGCAGCATACCGCGCTGAGTTCGAT 1487

RESULT 6
US-09-866-379-9

Sequence 9, Application US/09866379
Patent No. US20020136754A1

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
APPLICANT: O'DONOGHUE, Eileen
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIORITY FILING DATE: 2001-05-24
PRIORITY APPLICATION NUMBER: US 09/580,515
PRIORITY FILING DATE: 2000-05-25
PRIORITY APPLICATION NUMBER: US 09/318,528
PRIORITY FILING DATE: 1999-05-25
PRIORITY APPLICATION NUMBER: US 09/291,931
PRIORITY FILING DATE: 1999-04-13
PRIORITY APPLICATION NUMBER: US 09/259,214
PRIORITY FILING DATE: 1999-03-01
PRIORITY APPLICATION NUMBER: US 08/910,798
PRIORITY FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1

SEQ ID NO 9
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Query Match 97.28; Score 1286.2; DB 10; Length 1901;

Best Local Similarity 99.48; Pred. No. 0; Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCCGCAATCTGCA 60
188 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCCGCAATCTGCA 247
Qy 61 TTGCGTAGAGTGAGCGCGAGCTGCAAGCTGGAAGTGTGATTTGATGCTGATAGT 120
248 TTGCGTAGAGTGAGCGCGAGCTGCAAGCTGGAAGTGTGATTTGATGCTGATAGT 307
Qy 121 GTGGGTCTCCAAACCAAGGCGACCACTGATGAGATGTCACCCGAGCGATGAGCA 180
308 GTGGGTCTCCAAACCAAGGCGACCACTGATGAGATGTCACCCGAGCGATGAGCA 367
Qy 181 ACCTGGCGGTAAACTGAGTGGCTGACACCGCGNGTGTGATGAGCTAATCGCTATCTC 240
368 ACCTGGCGGTAAACTGAGTGGCTGACACCGCGNGTGTGATGAGCTAATCGCTATCTC 427
Qy 241 GGACATTACCAACGCCACGCTCTGTAGCCGAGGATTCGTCGCGAAAGGCGCGCG 300
428 GGACATTACCAACGCCACGCTCTGTAGCCGAGGATTCGTCGCGAAAGGCGCGCG 487
Qy 301 CAGTCTGAGTACGTCGCGATTAATGCTGATGTCGAGGAGGTACCGTTAAACAGCGAA 360
488 CAGTCTGAGTACGTCGCGATTAATGCTGATGTCGAGGAGGTACCGTTAAACAGCGAA 547
Qy 361 GCTTTCCGCGCGCGGCTGCGACCTGACTGTGCAATAACCGTATACCGAGCGAGATAG 420
548 GCTTTCCGCGCGCGGCTGCGACCTGACTGTGCAATAACCGTATACCGAGCGAGATAG 607
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTCTATAAACTGGCTTTCCCACTGAGTAAGCG 480
608 TCCAGTCCCGATCCGTTATTTATCTCTCTATAAACTGGCTTTCCCACTGAGTAAGCG 667
Qy 481 AAGCTGATGACGCGATCTCAGCAGGAGGAGGATGATGCTGATTTACCGGGCAT 540
668 AAGCTGATGACGCGATCTCAGCAGGAGGAGGATGATGCTGATTTACCGGGCAT 727
Qy 541 CGGCAAAAGCGGTTTCGCAACTGGAAGGCGGCTTAATTTCCGCAATCAACTTGTGC 600
728 CGGCAAAAGCGGTTTCGCAACTGGAAGGCGGCTTAATTTCCGCAATCAACTTGTGC 787
Qy 601 CTTAAAGTGGAAGAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATGGAATCTC 660
788 CTTAAAGTGGAAGAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACATGGAATCTC 847
Qy 661 AAGGTGAGCGCGCAATGCTCATTTAACCGGTGCTGTAAGCTCGCATCAATGCTGAGC 720
848 AAGGTGAGCGCGCAATGCTCATTTAACCGGTGCTGTAAGCTCGCATCAATGCTGAGC 907
Qy 721 GAGATATTTCTCTGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGAGAGATCAAC 780
908 GAGATATTTCTCTGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGAGAGATCAAC 967
Qy 781 GATTCACACCAAGTGGAACACCTTCTGTAAGTGTGATAACGCGCAATTTTATTTGCTCA 840
968 GATTCACACCAAGTGGAACACCTTCTGTAAGTGTGATAACGCGCAATTTTATTTGCTCA 1027
Qy 841 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCGTTATTTGATTTGATCAGCGAGCG 900
1028 CGCAGCGCAGAGTGGCCCGCAGCGCGCACCCGTTATTTGATTTGATCAGCGAGCG 1087
Qy 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGATGACATTACCACTTCAGTCTG 960
1088 TTGACGCCCATTCACCGCAAAAACAGCGGTATGATGACATTACCACTTCAGTCTG 1147
Qy 961 TTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1020
1148 TTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1207
Qy 1021 ACCTTTCCGGACACGATTAATCTGCAAAATCTGCGCGGCACTGAGAGCTCACTG 1080

Db 1208 ACCGTCGCCGTCAGCCGATTAACACGCCGCCAGGTGTGACTGTGTGAACGCTGG 1267
Qy 1081 CGTCGGCTAAGCATTAACAGCCAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTCAG 1140
Db 1268 CGTCGGCTAAGCATTAACAGCCAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTCAG 1337
Qy 1141 CAGATGCGTATTAACAGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1200
Db 1328 CAGATGCGTATTAACAGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1387
Qy 1201 CTGGCAGATGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGTTTACGCA 1260
Db 1388 CTGGCAGATGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGTTTACGCA 1447
Qy 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTGAT 1300
Db 1448 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTGAT 1487

RESULT 7

US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match 97.0%; Score 1283; DB 10; Length 1901;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1290; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATGAAGCAGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCGCAATTCGA 60
Db 188 ATGAAGCAGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCGCAATTCGA 247
Qy 61 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGTGATGTTCAGTCTGATGCT 120
Db 248 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGTGATGTTCAGTCTGATGCT 307
Qy 121 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGCAAGTGTCAACCCGCAAGCATGGCCA 180
Db 308 GTGCGTGTCTCAACCAAGGCGCAGCACTGATGCAAGTGTCAACCCGCAAGCATGGCCA 367

Qy 181 AACTGCCCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCATATCTC 240
Db 368 AACTGCCCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCATATCTC 427
Qy 241 GGACATTACCAACGCCAGCTCTGTGTACCGACGGAATTTGCTGGCGAAAAGGCTGCCCG 300
Db 428 GGACATTACCTGGGCGCAGCGCTGTGTACCGACGGAATTTGCTGGCGAAAAGGCTGCCCG 487
Qy 301 CAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 488 CAGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 547
Qy 361 GCTTCCCGCGCGGCTGGCGACCTGATGTCATTAACCGTATACCGAGCAGATACG 420
Db 548 GCTTCCCGCGCGGCTGGCGACCTGATGTCATTAACCGTATACCGAGCAGATACG 607
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGGCCAACTGATACGCG 480
Db 608 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGGCCAACTGATACGCG 667
Qy 481 AACTGACTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Db 668 AACTGACTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 727
Qy 541 CGGCAAAAGCGGCTTCCGCAACTGGAACGGGTGCTTATTTTCCGCAATCAACTGTGCG 600
Db 728 CGGCAAAAGCGGCTTCCGCAACTGGAACGGGTGCTTATTTTCCGCAATCAACTGTGCG 787
Qy 601 CTTAAACGTGAAACAGAGAGAGAAAGCTGTTCATTAACGAGCATTAACCATCGGAATC 660
Db 788 CTTAAACGTGAAACAGAGAGAGAAAGCTGTTCATTAACGAGCATTAACCATCGGAATC 847
Qy 661 AAGGTAGCGCGGCAATGCTCATTAACCGGTGCGGTGAAGCCCTGCAATCAATCTGACG 720
Db 848 AAGGTAGCGCGGCAATGCTCATTAACCGGTGCGGTGAAGCCCTGCAATCAATCTGACG 907
Qy 721 GAGATATTTTCTCCGCAACAGCAGAGGAATGCGGAGCGGCGGTGGGAAGATCAAC 780
Db 908 GAGATATTTTCTCCGCAACAGCAGAGGAATGCGGAGCGGCGGTGGGAAGATCAAC 967
Qy 781 GATTCAACACAGTGGAAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 840
Db 968 GATTCAACACAGTGGAAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 1027
Qy 841 CGCACGCCAGAGTTGGCCGCGAGCGGCCACCCCGTATTTGATTTGATCATGAGCAGCG 900
Db 1028 CGCACGCCAGAGTTGGCCGCGAGCGGCCACCCCGTATTTGATTTGATCATGAGCAGCG 1087
Qy 901 TTGACGCCCATCCACGCCAAACAGCGGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1088 TTGACGCCCATCCACGCCAAACAGCGGATGATGATGATGATGATGATGATGATGATGAT 1147
Qy 961 TTTATTCGCGGACACGATTAATCTGCAAAATCTCGCGCGCGGCGACTGAGACTCAACTGG 1020
Db 1148 TTTATTCGCGGACACGATTAATCTGCAAAATCTCGCGCGCGGCGACTGAGACTCAACTGG 1207
Qy 1021 ACGTTCGCCGCTGACGCCGAGTAAACAGCCCGCGAGGTGTAACCTGAGTGTGTAACGCTGG 1080
Db 1208 ACGTTCGCCGCTGACGCCGAGTAAACAGCCCGCGAGGTGTAACCTGAGTGTGTAACGCTGG 1267
Qy 1081 CGTCGGCTAAGCATTAACAGCCAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTCAG 1140
Db 1268 CGTCGGCTAAGCATTAACAGCCAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTCAG 1337
Qy 1141 CAGATGCGTATTAACAGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1200
Db 1328 CAGATGCGTATTAACAGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1387
Qy 1201 CTGGCAGATGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGTTTACGCA 1260
Db 1388 CTGGCAGATGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGTTTACGCA 1447
Qy 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTGAT 1300

Db 1448 ATCGTGAATGAAGCAGCATACCGCGGTGCTGATTTGTAAT 1487

RESULT 8

US-10-266-041-9
 ; Sequence 9, Application US/10266041
 ; Publication No.: US20030072844A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Iel, Xingen
 ; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
 ; FILE REFERENCE: 19603/2791
 ; CURRENT APPLICATION NUMBER: US/10/266, 041
 ; CURRENT FILING DATE: 2002-10-07
 ; PRIOR APPLICATION NUMBER: US/09/540, 149
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: 60/127, 032
 ; PRIOR FILING DATE: 1999-03-31
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 9
 ; LENGTH: 1489
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 US-10-266-041-9

Query Match 95.5%; Score 1263.8; DB 9; Length 1489;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1280; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAACCCCGCAATCTGCA 60
 DB 182 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAACCCCGCAATCTGCA 241
 QY 61 TTGCGTCAGATGAGCGGAGCTGAAGCTGAAAGTGTGATTTGTCAGTCGTCATGT 120
 DB 242 TTGCGTCAGATGAGCGGAGCTGAAGCTGAAAGTGTGATTTGTCAGTCGTCATGT 301
 QY 121 GTGCGTGTCTCAACCAAGCCAGCAGCACTGATGAGATGTCAACCCAGAGCGATGCCA 180
 DB 302 GTGCGTGTCTCAACCAAGCCAGCAGCACTGATGAGATGTCAACCCAGAGCGATGCCA 361
 QY 181 ACCTGGCGGTTAAACGAGGTGGTGGTGCACCGGNGGTGGTGAATTCGCTATCTC 240
 DB 362 ACCTGGCGGTTAAACGAGGTGGTGGTGCACCGGNGGTGGTGAATTCGCTATCTC 421
 QY 241 GGACATTAACCAAGCCAGCTGTGTAGCCGAGCGATTTGTCGCGAAAGGCGTCCCG 300
 DB 422 GGACATTAACCAAGCCAGCTGTGTAGCCGAGCGATTTGTCGCGAAAGGCGTCCCG 481
 QY 301 CAGCTGTGTAGCTGTGATTTATTTGCTGATGTGAGAGCGATVACCGTTAAACAGCGAA 360
 DB 482 CAGCTGTGTAGCTGTGATTTATTTGCTGATGTGAGAGCGATVACCGTTAAACAGCGAA 541
 QY 361 GCCTTGGCGCGCGGCTGGAGCGACTGTGCAATACCGTACATACCGAGCGATAGC 420
 DB 542 GCCTTGGCGCGCGGCTGGAGCGACTGTGCAATACCGTACATACCGAGCGATAGC 601
 QY 421 TCCAGTCCGATCGTTATTTATCTCTTAAAACTGGCGTTTGCACCACTGGATAGCG 480
 DB 602 TCCAGTCCGATCGTTATTTATCTCTTAAAACTGGCGTTTGCACCACTGGATAGCG 661
 QY 481 AACGTGACTAGCGGATCTCAGCAGGAGGAGGAGGCTCAATTGCTGACTTACCGGCAT 540
 DB 662 AACGTGACTAGCGGATCTCAGCAGGAGGAGGAGGCTCAATTGCTGACTTACCGGCAT 721
 QY 541 CGGCAAAAGCGGTTTGGCGAACTGGAACGGGTCTTAATTTTCCGCAATCAATTTGTGC 600
 DB 722 CGGCAAAAGCGGTTTGGCGAACTGGAACGGGTCTTAATTTTCCCAATTAACCTTTGTC 781
 QY 601 CTTAAAGCTGAAAGAGAGAAAGCTGTTCATTACGAGGATTAACCATTAACGGAATGC 660
 DB 782 CTTAAAGCTGAAAGAGAGAAAGCTGTTCATTACGAGGATTAACCATTAACGGAATGC 841

QY 661 AAGGTAGCGCGAGCAATGTCTCATTAACCGGTGCGGTAACCTCGCATCAATGCTGAGC 720
 DB 842 AAGGTAGCGCGAGCAATGTCTCATTAACCGGTGCGGTAACCTCGCATCAATGCTGAGC 901
 QY 721 GAGATATTTCTCTGGAACAAGCAGAGGAAATGCGGAGCGGGGTGGGGAAGATCAC 780
 DB 902 GAGATATTTCTCTGGAACAAGCAGAGGAAATGCGGAGCGGGGTGGGGAAGATCAC 961
 QY 781 GATTACACCAAGTGAACACCTTGCTAAGTTTGCATTAACCGCAATTTATTTGTCACAA 840
 DB 962 GATTACACCAAGTGAACACCTTGCTAAGTTTGCATTAACCGCAATTTATTTGTCACAA 1021
 QY 841 CGCAGCGCAAGTGTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 DB 1022 CGCAGCGCAAGTGTGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1081
 QY 901 TTGACGCGCCATCCAGCGCAAAAGAGCGGTATGTGATTAACCACTTCAGTACTG 960
 DB 1082 TTGACGCGCCATCCAGCGCAAAAGAGCGGTATGTGATTAACCACTTCAGTACTG 1141
 QY 961 TTTATTGCGGAGACGATTAATCTGCAAAATCTGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 DB 1142 TTTATTGCGGAGACGATTAATCTGCAAAATCTGCGGCGGCGGCGGCGGCGGCGGCGG 1201
 QY 1021 ACCTTCCCGGTACGCGGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 DB 1202 ACCTTCCCGGTACGCGGATTAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1261
 QY 1081 CGTCGCGTAAAGCATAACAGCCAGTGTGATTCAGTTTCGCTGCTTCAGACTTTACAG 1140
 DB 1262 CGTCGCGTAAAGCATAACAGCCAGTGTGATTCAGTTTCGCTGCTTCAGACTTTACAG 1321
 QY 1141 CAGATGCGTAAATAAACCGCGCTGTCTTAATTAACGCGCGGCGGCGGCGGCGGCGGCGG 1200
 DB 1322 CAGATGCGTAAATAAACCGCGCTGTCTTAATTAACGCGCGGCGGCGGCGGCGGCGGCGG 1381
 QY 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
 DB 1382 CTGGCAGAGTGTGAAGAGCGAAATGCGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1441
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAAGTTTGATTCATCAC 1308
 DB 1442 ATCGTGAATGAAGCAGCATACCGCGCTGCAAGTTTGATTCATCAC 1489

RESULT 9

US-10-021-723A-3
 ; Sequence 3, Application US/10021723A
 ; Publication No.: US20030101476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Short, Jay
 ; APPLICANT: Mathur, Eric
 ; APPLICANT: Richardson, Toby
 ; APPLICANT: Robertson, Dan
 ; APPLICANT: Barton, Nelson
 ; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
 ; FILE REFERENCE: 112766.140 (Div-016CIP)
 ; CURRENT APPLICATION NUMBER: US/10/021, 723A
 ; CURRENT FILING DATE: 2002-10-23
 ; PRIOR APPLICATION NUMBER: US 60/255, 090
 ; PRIOR FILING DATE: 2000-12-12
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 1325
 ; TYPE: DNA
 ; ORGANISM: Yersinia pestis
 US-10-021-723A-3

Query Match 16.5%; Score 218.8; DB 9; Length 1325;
 Best Local Similarity 52.8%; Pred. No. 9e-67;
 Matches 567; Conservative 0; Mismatches 493; Indels 13; Gaps 4;

```

RESULT 10
US-10-021-723A-1
; Sequence 1, Application US/10021723A
; Publication No. US20030101476a1
; GENERAL INFORMATION:
; APPLICANT: Mathur, Eric
; APPLICANT: Short, Jay
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (Div.-016C1P)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1325
; TYPE: DNA
; ORGANISM: Yersinia pestis
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 470, 487, 491, 492, 522
; OTHER INFORMATION: n - A,T,C or G
US-10-021-723A-1

Query Match          16.3%; Score 215.6; DB 9; Length 1325;
Best Local Similarity 51.9%; Pred. No. 1,2e-65;
Matches 556; Conservative 0; Mismatches 505; Indels 11; Gaps 3

89  TGAAGAAGTGTGATGTGTCAGTGCCTCAATGATGTGTCGTCCTCAACCAAGGCCACGCAAC 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
107 TAGAAGCGTGGTTATTTTGAAGTCCCATGTTGCGTCGCGCAACCAACCAAGGCCAC 166

149  TGAATGACGATGTCAACCCACAGCAGCATGGCCCAACTGGCCGTTAAACTGGGTTGGCTCA 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 TTATGAAAGATGTTACGCCAGATAGTGGCCCAATGGCCGGTTAAAGCGGGGTATTTTAA 226

209  CACGCCGNGGTGTGATAGCTAATCCGCTATCTCGGACATTACCAAGCCAGCGCTGTGGTAG 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 CGCCACGCGTGTGCGAGTGTGTCACATGTATGATGGGGGGTTTATGTGATTTACTTTCCGA 286

269  CCGACGATGCTGTGCGAAAAAGGGCTGCCGCACTGTGTGAGTCTGCGGATTAATTTCTG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
287 GCCTGTGTTGTATACG--GCGGGATGTGCCGACAGAGGGGGGTATATGCACAGCGAG 343

329  ATGTGACGAGCGTACCCGTTAAAGAGGCGAAGCCCTGCGCCCGCGGCGTGGACCTGACT 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
344 ATATGATCAACGTACCCGCTTAAACCGGACAGGCAATTTCTTATGTGTGGCTCCGGGGT 403

389  GTGCATTAACCGTATACCCAGGCGAGATACGTCCAGTCCCGATCCGTTATTTAATCTCTC 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 GTGGTTTGACCGTCAATATAGCGCCGATTTGAAAAAGACGATCCCGCTTCATCCGG 463

449  TAAAAACGTGGCGTTTGCCCACTGGATTAACGCGAAGCTGACTACGCGATCTCTCAGAGGG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
464 TAGAGAGTGGCGTGTGTAAAGTTAANCGNNCGCAAAACAGATTAACGATTAAGAAACANT 523

509  CAGAGAGGTCAATTCCTGACTTTACCGGGCATCGGCAAAAGGCGTTTGGCGAAGTGAAC 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
524 TGGCGGCGCGCTTAATACGGGTGAGCGACGCGTACGCTTAACCTTTTGGCCAGATGGGG 583

569  GGGTGTCTAATTTTCCGCAATCAAACTGTGCTTTAAAGTGAGAAACAGGCAAGAACT 628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
584 ACGTGTCAATTTTGGCGGCTTCTCTTATTCGAATCTTTGCAACAGCAAGAAAAAAGCT 643

629  GTTCATTACGACGACATTACCATTCGGAACATCAAGTGAAGCGCCA-----CAATGTCTC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
644 GTGATTTTGGCACTTTGGCGCCAAATGAAGTTAACGTTAATTAAGAAGGACCAAAAGTGC 703

```

```

QY 684 ATTACCGGTGCGGTAAGCTCCGATCATGCTGACGAGATATTTCTCGACAAC 743
Db 704 CCTCAGTGGGCGACATGCGTATCATGACATGCGTAATCTCTTAATTAACAAACG 763
QY 744 ACAGGAAATGCGGAGCGGCGGAGGAGATCAGATTCACACAGAGGAACACCTT 803
Db 764 ACAGGCAATGCGGAGGTTGCTGCGCAACGCTTAAAGGGGCGAGATGCGATCTT 823
QY 804 GCTAAGTTTCATACGCGCAATTTTATTTGCTACAAACGACGCGCAGAGTTGCCCGAG 863
Db 824 ATTGTCTATCATATACGCGCAATTTAATTTAATGCAAAAACCGATATATCGCCGTCA 883
QY 864 CCGGCGCACCCCGTATTTGATTTGATCATGCGCTTACGCCCATCATCAC--CGCA 920
Db 884 TAAAGGACCGCATATTTACAGCAGATGATACGCTTAAACCTTCACCTGATGCTCA 943
QY 921 AAAACAGCGGTATGCTGACATTTACCATTCAGTACTGTTATTTCCGCGACAGATAC 980
Db 944 GGGGCAAAACCTACCCATTTTCAGCCCAAAACCGGCTTGTCTCGGTGGGCAATGATAC 1003
QY 981 TAACTGCGCAATCTCGGCGCGCATGAGCTCACTGACGCTTCCCGTCAGCCGGA 1040
Db 1004 CAATATGCGCAATATGCTGGGTATGCTGGAGCGCATGGAGCTACCCGACGACCTGA 1063
QY 1041 TAACAGCGCGCGAGTGTGTAATGCTTTGAACGCTGCGCTGAGCTAAGCATAACAG 1100
Db 1064 TAATATCTCAACAGTGTGGGATGTTTGAACATATGCGCAAAATCCGATTAACACCA 1123
QY 1101 CCAATGATTCAGATTTGCTGCTGCTCCAGACTTTCAGACAGATGCGGTAT 1152
Db 1124 GCGCTACGTTGCGGTGAAGATGTTCTACCAACGATGATCAGTTGCGTAAAT 1175

```

```

RESULT 11
US-10-021-723A-11
; Sequence 11, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021.723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Yersinia pestis
US-10-021-723A-11

```

```

Query Match 15.6%; Score 206.6; DB 9; Length 1326;
Best Local Similarity 51.0%; Pred. No. 2e-62;
Matches 569; Conservative 0; Mismatches 335; Indels 12; Gaps 3;

```

```

QY 77 CGGAGCTGAAGCTGGAAAGTGTGATTTGCTACGTCTCATGTCGCTGCTCAACCA 136
Db 95 CCGGCTACACCTTAAGAACTGTAGTATTTTGAAGTCCCATGCTGCTGAGCCGACTA 154
QY 137 AGGCGACGCACTGATGATGATGCTACCCGACGATGCGCAACCTGCGCGTAAAC 196
Db 155 AGCAGACGCGAGCTGATGATGATGATACACCTGTAAGTGGCTCAGTGGCGGATTAAG 214
QY 197 TGGGTTGGCTGACACCGCGAGTGTGAGCTAATCGCTATCTCGGACATTAACCAAGCC 256
Db 215 CCGGCTATTGACCTCCGTCGTGGCGCGCACTGTCACACCTGATGAGGGGCTTATGCGG 274

```

```

QY 257 AGCGTCTGTAACCGGACGATTTGCTGCGCAAAAAGGCTGCCGCGACTGTGTCAGTCG 316
Db 275 ATTAATTTCCCGAGTTTGGCTTTTGGC---CGGGGCTGCGGAGAGGGGCGGTAT 331
QY 317 CGATTTATGCGATGCGAGAGGCTACCGCTTAAACAGGCGAAGGCTTCCGCGCGGGC 376
Db 332 ATGCAACAGGAGATATGACACGCGCATGCTTTTAAACGGTCAAGCTTTTCTGATG 391
QY 377 TGGCAGCTGACCTGTGCAATTAACCGTATACCCAGCGAGATAGCTCCAGTCCGATCCG 436
Db 392 TGGCGCGCGGGTGGCGCTGACTGCTCCACATGAGGCGGATCTTAAGAAAACGATCTC 451
QY 437 TATTTATCTCTTAAACCTGCGCTTTGCCAATCTGATTAACGCGAAGCTGATGACCGGA 496
Db 452 TGTTCATATCCCGTTGAACACCGGCGCTGTGAATGCGCAACCGCCAAACCGATTAAGCAA 511
QY 497 TCCCTACAGAGGCGAGAGGCTCAATTCCTGACTTACCGGCGCATCGCAACGCGCTTC 556
Db 512 TTGAGGAACGCTGCGGCGCGCGCTTACAGACAGGTAACCGCATGCGCTATGCCAAACGCTTTG 571
QY 557 GCGAATGGAACGCGGCTCTTAATTTCCGCAATCAAACTTGCTTAAACGTGAGAAAC 616
Db 572 CGCAATATGGGAGATGCTCTGATGACTGCTGAGATCCGTAAGTCACTGAGAGAGC 631
QY 617 AGCAGCAACGCTGTTCAATTAACGAGCATTTACATGGAATCAAGTACAGCGCGCAC- 675
Db 632 AGGGAAACCTGTGACTTGCACACTTTCGCGGCAAGATTAATGTAACCAAGGAAG 691
QY 676 -----AATGTCATTAATTAACCGGTGCGGTAAAGCTCGCATCAATGCTACGAGATATTC 730
Db 692 GCACGAAAGTTACCTGTACAGGCCCTGCGCGCTGTCTAGCAGCTTGGGGAATCTTC 751
QY 731 TCCTGCAACAAGCAGCAGGATGCGGAGCGCGGAGGAGATCAGGATTCACACC 790
Db 752 TGCTGCAAGACGCGCAGCGCATGCGCGAAGTAGGTGGAGCGCTTGAAGGCGCTGAGA 811
QY 791 AGTGAACACCTTCTGATGTTGATTAACGCGCAATTTTATTTGCTACAAACGCGCGAC 850
Db 812 ACTGGGTGCTCTTCTGAGCTGCAATGCTCAACGTTCAACCTGATGCGTAAACGCGCAT 871
QY 851 AGGTGCGCGGAGCGCGGCGCACCGCTTATTTGATTTGATGATGAGCGCTTGAACGCGCC 910
Db 872 ACATTTGACCGCCCAAAAGCGAGCGCTTTTACGCAAAATCGATACCGCATGACCTGCG 931
QY 911 ATCCAC---CGCAAAAACAGCGGATGCTGATGATTAACCACTTCACTAGTCTTATTTG 967
Db 932 AACTGAGACGCGCAGGCGCAAAAACCTGCGATCTGCGCTCAGAAACCGGTTTATTCCTG 991
QY 968 CCGGACACGATTAATCTGCAAAATCTCGCGCGCGCATGAGCTCACTGAGAGCTTC 1027
Db 992 GTGGCCACGACCAAAATTTGCTTAACATCCCGGTATGCTGGGCGAGATTGGCGATTAC 1051
QY 1028 CCGGTGACCGGATTAACGCGCGCGCGAGTGTGATGCTGTTGAACGCTGCGGTGGG 1087
Db 1052 CGGAAACACCGGATTAACACCCACCGCGCGGTCTGCTTTGAGCTGTGCGAGATTC 1111
QY 1088 TAAACGATTAACAGCGATGATTCAGTTTCGCTGCTTCCAGACTTTAACAGCAGATC 1147
Db 1112 CGGACATATCAATCAAGTTATGTCGCTTAAGATGTTCTATACAGACCATGATCAATTCG 1171
QY 1148 GTGATTAACAGCGGCTGCTCAATTAATACGCGCGCGC 1183
Db 1172 GTAACGCGGAGAGCTGATTTAAAGAAATCCG 1207

```

```

RESULT 12
US-10-021-723A-9
; Sequence 9, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby

```

```

; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 11276.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 9
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 954-2 phytase sequence
; US-10-021-723A-9

```

```

Query Match      12.1%; Score 159.6; DB 9; Length 1431;
Best Local Similarity 49.3%; Pred. No. 1,1e-45;
Matches 575; Conservative 0; Mismatches 565; Indels 27; Gaps 5;

```

```

QY 75 GCGGAGCTGAGTGAAGTGTGATGTCAGTGTGATGTCAGTGTGTCGTCGTCAC 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 168 GCAAGACTGTGTCGTCGACGCGTGTGTCATGTCAGCGCGGACGCGGTCGCGAC 227
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 135 CAAGGCGACCACTATGAGATGTCACCCAGACGCGATGGCCAACTGGCCGCTAA 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GCGTCCGAGCCCGCTGAGAGCCCTGACGCGCCGATCCGTCGCGCCGCGCGGCC 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 ACTGGGTTGGCTGACACACCGCGGTGAGCTAATGCTATCTGACATTTACAAAG 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 288 GATGCGCCACTGACGATGTCGCGCGGCGCTGTCTGCGAGATGGGGGCTACTACG 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 CCAAGCTGTGTGATGCGAGGATGTCGCGGAAAAGGCTGCCGAGTCTGTGTCAGT 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 TGATTTGCTTCGTCGCGCGGCGGTGTGTCGCGCGACGCGGCTCTGACCGGAGCT 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 CCGGATTTGCTGATGTCGACGACGCTACCCGTAACGAGCGCAACCTTCGCGCGG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 408 TTACGATGCGACGATGTCACGACGCGCGCTGACGCGCGACGCGCTGCTCTCGG 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 375 GCTGCGACCTGATGTCATTAACCTACATACCGACGAGATACGTCCGATCCGATCC 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 468 CATGGCGCCGAGGCTGCTGATCCACAGCATCATCGGCGCGCTGACAGAGAGATCC 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 435 GTTATTTAATCTCTTAAACCTGCGCTTTCGCACTGATTAACCGAAGCTGACTGAC 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 528 GATCTTCACGCGATGGAATCGGTCGATGCCAGTCGACCCCTACAGGGAAGCGCGA 587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 495 GATCTTCACGAGGCGAGAGGATGTCGATTTACGGCGCATGGCAACGGCGTT 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 588 CATCGAAGCGCATGCGCGGAG--GCGGCGGTGGCGACACTGGGAAAGCGCTAGCGACA 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 555 TCGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGCTTAAACGTGAGAA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 646 AGCGTGACCAAGATGAGCGAGGTCTGATACGCCATAGCGCGGATTTGGCCAGCAT 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 615 ACAGGAGCAAAAGCTGTTCATTAACGAGCATTAACATCGAATCAAGGTAGCGCGA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 706 GCGGCGCAATGCGATGTCGCGCGCAACCAATGTCGAGATCAGACAGATGGCTT 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 675 CAATGTCTCATTAACCGGTGCGGTAGCTTCGATCATATGTCAGGAGATATTTCCT 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 766 C-ATGCGCGCTTGAAGGCGCGATGAGCTTCACGCTCCAGAGCTTCCTGCT 824
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 735 GCAAGACGACAGGATGCC-----GAGCGCGGGTGGGAGAGATACGACATTACA 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 CGAAGCATGGGCGAGGCTGCGCACAGAGCAGGTTCATGGGCGCTTACAGATGGCGCA 884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 789 CAGGTGAGACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAAACGACGCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 GGAAGTGAAGCTGCTGTATGACAGGCGCATTAACGCGCAGTGTGATTTGGGGAAGAGCGCC 944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

QY 849 AGAGTTGCCCGAGCGCGCCACCCGTTATTGATTTGATCATGACGAGT----- 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 TTACATGCTACTGCAAGAGGCGACCGGATGCTGCTTCGCTGATGCGCTTGACGG 1004
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 902 -----TGACGCCCATCATCAGCCGCAAAAGCGGTATGATGATTAACCACTTAC 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1005 GCGGCGTGGGCGCCACCTCGGAGCTTGGCGTTCAAGGCGCCAGAGTGGCCCAAGGCA 1064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 956 TACTGTTTA-----TTCCCGACACGATTAATCTGCGCAATCTGCGCGCGCATGGA 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1065 CCGGTCTATGTGTGACCGCGCATGACAGAACTTTCACACTTGGCCGCTTGCGCA 1124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1011 GCTCACTGACGCGTCCCGGTGCGGATTAACAGCCGCGAGGTGATGATGATGAT 1070
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 CTGCGATGACGCTGCGCCGACGACAGCAGACGCGCGCGCGCTGATGATGAT 1184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1071 TGACGCTGCGCTGCGCTAAGCGATTAACAGCCAGTGTGATGATGATGATGAT 1130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 CTCTGTTGGGCGGAGCGCTGCGACGAGGACGTTGCTGCTGCGGATGATGATGAT 1244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1131 GACTTTACAGCAGATGCTGATTAACAGCCGCTGCTATTAATACGCGCGCGAGAGGT 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1245 GTCGATGATGATGCTGCGGAGCTACGCGCTCTCCCTGCGCGAGCGCGCATGCT 1304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1191 GAACCTGACCCCTGCGAGATGTGA 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1305 GATCTTCCGCTGCGCGCTGTCGCA 1331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-021-723A-7
; Sequence 7, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 11276.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1230
; ORGANISM: Rhizobium
; US-10-021-723A-7

Query Match      5.5%; Score 73; DB 9; Length 1230;
Best Local Similarity 54.1%; Pred. No. 7.1e-15;
Matches 170; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

```

QY 321 TATTGTGATGTGACGAGCGTACCCCTAAACAGCGGAAGCTTTCGCCCGCGGCTGCC 380
DB 330 CAAGGCGAGCTACAAAGAGCGGACGATGCCACCGCGAGAGCTGGCGCGGGGCTTCAT 389
QY 381 ACCTGACTGTGCCA 394
DB 390 GCCCGGCTGGCGCA 403

RESULT 14

US-10-021-723A-5

Sequence 5, Application US/10021723A
Publication No. US20030101476A1
GENERAL INFORMATION:
APPLICANT: Mathur, Jay
APPLICANT: Short, Eric
APPLICANT: Richardson, Toby
APPLICANT: Robertson, Dan
APPLICANT: Barton, Nelson
TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
FILE REFERENCE: 112766.140 (DIV-016CIP)
CURRENT APPLICATION NUMBER: US/10/021,723A
PRIORITY FILING DATE: 2002-10-23
PRIORITY APPLICATION NUMBER: US 60/255,090
PRIORITY FILING DATE: 2000-12-12
NUMBER OF SEQ ID NOS: 74
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1266
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 953-6 phytase sequence
US-10-021-723A-5.

Query Match

4.9% Score 64.8; DB 9; Length 1266;

Best Local Similarity 52.88; Pred. No. 6e-12; Mismatches 143; Indels 3; Gaps 1;

Matches 163; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 81 GCTGAACGTAAGAGTGTGATGTCAGTCGTCATGAGTGTCGTCCTCCCAACCAAGC 140
DB 129 GCTAACTGTAAGAGTGTGATGTCAGTCGTCATGAGTGTCGTCCTCCCAACCAAGC 188
QY 141 CACGCAATGATGTCAGTGTGTCACCCAGACGATGCGCAACCTGGCGGTAACCTGG 200
DB 189 GCGGCTGTGTCGCGCGCGGTTATTCGCGCAACATGCGCGGATGCGGATTCGCG 248
QY 201 TTGGCTACACCGGNGCTGTGATGTCATGCTGCTATCTGGACATTACCAAGCCAGC 260
DB 249 CTGCTGACCGCGGACGCGGCGGCGGCTCAAGCTGCTGCGCAAGCGCGCTCAC 308
QY 261 TCTGTGACCGGACGATGTCGCGCAAGAGGCTGCGCGAGTCTGTGATGTCGCGAT 320
DB 309 TTTCGGGCTGCGCGCTATTCGCG---GACGCTGTCGCGCGCGCGGACGATGCTCT 365
QY 321 TATTGTGATGTGACGAGCGTACCCGTAACAAAGCGGAAGCTTTCGCCCGGCTGCC 380
DB 366 CAAGGCGAGCTACAAAGAGCGGACGATGCGAGCGCGCAAGATGCGCGGCTTCAT 425
QY 381 ACCTGACTG 389
DB 426 GCCCGGCTG 434

RESULT 15

US-10-184-644-312

Sequence 312, Application US/10184644
Publication No. US20030044930A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Matanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C227
CURRENT APPLICATION NUMBER: US/10/184,644
PRIORITY FILING DATE: 2002-06-28
PRIORITY APPLICATION NUMBER: See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 312
LENGTH: 802
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-644-312

Query Match

2.8% Score 36.4; DB 9; Length 802;

Best Local Similarity 5.2%; Pred. No. 0.06; Mismatches 32; Conservative 190; Mismatches 395; Indels 0; Gaps 0;

QY 682 TCATTAAACGGTGGTGAAGCTTCGATCAATGCTGACGAGATATTCTCTGCAACA 741
DB 112 TTPKVOFFSESGSDTVPVPLRGVDSYPPQKSFMLKMHYLDKYEWMRAD 171
QY 742 GCACAGGGAATGCCGCGGCGGCGGAGGAGATACCGATTCACACCGAGGACAC 801
DB 172 DDVTKGDRLENFRLNSSEPLFLGOTGLTTEEMKRLALPEBENCMGPGVIMREV 231
QY 802 TTGCTAAGTTTGATACGCGCAATTTTATTTGCTACACGACCGGAGGTTGCCG 861
DB 232 LRNVPHIKRCLRMVTHDEVEGRVRRPAGVQVWSEMRQLFEENEQNKGYRD 291
QY 862 AGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 921
DB 292 LHSKIQHAIYTLHKNKMPPOYRLHSYMLSKISLHRTIQLHREIVLSKSYNTEIRK 351
QY 922 AAACAGGCGATGATGTCGATTCACCATTCAGTACTGTTTATTCGCGGACGATCT 981
DB 352 EDIOLGTPSPFMRQPPORREILEMEFLTKLYLSAVDGPFRGMDSAOREALDLYM 411
QY 982 AATCTGCAATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1041
DB 412 VMEMINMANTRGRILDFKEIYGRVNMVYGAETLIDLLLYKKHKKMTPVVRHA 471
QY 1042 AACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1101
DB 472 YLAOTPFKIOFVEHEBELDAOELAKRINQESGSLFSLNSLTKLVFPQLPSKSEHKRPK 531
QY 1102 CAGTGGATGAGGTTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1161
DB 532 KKNILIPLSGRFMEVRFMGNFEKTLIPNQNKLVLFLPNSDNDKAKOVELMDYR 591
QY 1162 CTGTCAATTAATACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1221
DB 592 IKYPKADMQLPLPVSERSALALEVSSQFNNEBLFFCVDVLFTEFIFORCANVVLG 651
QY 1222 AATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1281
DB 652 QOIYFPIIFSOYDPKIVYSGKVPDSNHFATOKTGFWRNYFGITCIYKGLDVRVGGFDV 711
QY 1282 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1298
DB 712 SIQGWGLEVDVLEKNVY 728

Search completed: June 12, 2003, 10:27:21
Job time : 244 secs

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

OM nucleic - nucleic search, using sw model

(without alignments)
10815.453 Million cell updates/sec

Perfect score: 1323

Sequence: 1 atgaaagcatcttaatccc.....atcacatcacatcactaa 1323

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

```
Minimum DB seq length: 0
```

Maximum DB seq length: 200000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

GenEmbl: *

1: gb_ba:*

2: gb_hhg:*

3: gb_ln:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_scs:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_ln:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_scs:*

28: em_un:*

29: em_vl:*

30: em_hhg_hum:*

31: em_hhg_ln:*

32: em_hhg_other:*

33: em_hhg_mus:*

34: em_hhg_pln:*

35: em_hhg_rod:*

36: em_hhg_nam:*

37: em_hhg_vrt:*

38: em_sy:*

39: em_hhg_hum:*

40: em_hhg_mus:*

41: em_hhg_other:*

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	1322	99.9	1323	6	AR108133	AR108133 Sequence
2	1322	99.9	1323	6	AR127818	AR127818 Sequence
3	1322	99.9	1323	6	AR130956	AR130956 Sequence
4	1322	99.9	1323	6	AX052713	AX052713 Sequence
5	1322	99.9	1323	6	AX338966	AX338966 Sequence
6	1322	99.9	1323	6	AX356566	AX356566 Sequence
7	1287.8	97.3	1901	1	ECOAEP72	MS8708 Escherichia
8	1287.8	97.3	1901	6	AX356572	AX356572 Sequence
9	1287.8	97.3	10784	1	AE000200	AE000200 Escherich
10	1287.8	97.3	18882	1	D90735	D90735 Escherichia
11	1286.2	97.2	1901	1	ECOAEPHB	LO3371 Escherichia
12	1286.2	97.2	1901	6	AX356570	AX356570 Sequence
13	1286.2	97.2	1901	6	AX356574	AX356574 Sequence
14	1284.6	97.1	1901	1	ECOAEPHC	LO3372 Escherichia
15	1284.6	97.1	1901	1	ECOAEPHD	LO3373 Escherichia
16	1284.6	97.1	1901	1	ECOAEPHE	LO3375 Escherichia
17	1284.6	97.1	1901	1	ECOAEPHF	LO3376 Escherichia
18	1283	97.0	1901	1	ECOAEPHA	LO3375 Escherichia
19	1283	97.0	1901	6	AX356571	AX356571 Sequence
20	1282.2	96.9	3470	6	AX042376	AX042376 Sequence
21	1282.2	96.9	4060	6	AX042374	AX042374 Sequence
22	1282.2	96.9	6116	6	AX042375	AX042375 Sequence
23	1282.2	96.9	6708	6	AX042373	AX042373 Sequence
24	1282.2	96.9	17732	6	AX042378	AX042378 Sequence
25	1282.2	96.9	20623	6	AX042372	AX042372 Sequence
26	1270.2	96.0	5421	6	AX042377	AX042377 Sequence
27	1246.6	94.2	10029	1	AE005292	AE005292 Escherich
28	1246.6	94.2	327773	1	AP002554	AP002554 Escherich
29	713.4	53.9	212936	2	AC020970	AC020970 Mus muscu
30	319.6	24.2	114979	2	AC020798	AC020798 Mus muscu
31	305.8	23.1	700	1	BCAPPA	X05471 E.coli appa
32	232.8	16.8	10029	1	AE013783	AE013783 Yersinia
33	222.8	16.8	193050	1	AJ041419	AJ041419 Yersinia
34	160.2	12.1	178373	2	AC102243	AC102243 Mus muscu
35	97.8	7.4	825	1	AF427147	AF427147 Raoultell
36	65.6	5.0	11877	1	AE011683	AE011683 Xanthomon
37	64.8	4.9	21670	1	AE008748	AE008748 Salmonell
38	62.4	4.7	254050	1	AL627269	AL627269 Salmonell
39	61.6	4.7	10345	1	AE005727	AE005727 Caulobact
40	61.4	4.6	10585	1	AE011713	AE011713 Xanthomon
41	60.6	4.6	5166	1	STU75949	U75949 Salmonella
42	56.8	4.3	1675	1	ECOA6PA	M3807 E.coli peri
43	56.8	4.3	11710	1	AE000202	AE000202 Escherich
44	56.8	4.3	18252	1	D90737	D90737 Escherichia
45	55.4	4.2	13498	1	AE012175	AE012175 Xanthomon

ALIGNMENTS

RESULT 1	ARI08133	1323 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	ARI08133				
DEFINITION	Sequence 1 from patent US 6110719.				
ACCESSION	ARI08133				
VERSION	ARI08133.1	GI:12823620			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1323)				
TITLE	Kretz, K.				
JOURNAL	Phytase				
FEATURES	Patent: US 6110719-A 1 29-AUG-2000;				
	Location/Qualifiers				

source	1..1323	/organism="unknown"	
BASE COUNT	323 a	353 c	357 g
ORIGIN			289 t
			1 others
Query Match	99.9%;	Score 1322;	DB 6;
Best Local Similarity	100.0%;	Pred. No. 0;	Length 1323;
Matches 1323;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1	ATGAAGCGATCTTAATCCATTTTAACTCTTCTGATTCGGTTAAACCCGCAATCTGCA	60
DB	1	ATGAAGCGATCTTAATCCATTTTAACTCTTCTGATTCGGTTAAACCCGCAATCTGCA	60
OY	61	TTGCGTCAGATGAGCGGAGCTGAAAGCTGGAAATGGTGGTATGTCAAGTCGTATGT	120
DB	61	TTGCGTCAGATGAGCGGAGCTGAAAGCTGGAAATGGTGGTATGTCAAGTCGTATGT	120
OY	121	GTGCGTCTCCAAACCAAGCCAGCAACATGATAGAGATGTCACCCAGACATAGCGCA	180
DB	121	GTGCGTCTCCAAACCAAGCCAGCAACATGATAGAGATGTCACCCAGACATAGCGCA	180
OY	181	ACCTGCGCGGTAAACTGGTTGGCTGACACCCGCGNGTGGTGGAGCTAATCGCTATCTC	240
DB	181	ACCTGCGCGGTAAACTGGTTGGCTGACACCCGCGNGTGGTGGAGCTAATCGCTATCTC	240
OY	241	GGACATTTACCAAGCCAGCGCTGTGTAGCCAGAGATTGTCTGCGCAAAAGGCGTCCG	300
DB	241	GGACATTTACCAAGCCAGCGCTGTGTAGCCAGAGATTGTCTGCGCAAAAGGCGTCCG	300
OY	301	CAGTCTGTCAGTGCAGCTATTATTTGCTGATGTGCAAGAGCGTACCCTGTAACACGCGNA	360
DB	301	CAGTCTGTCAGTGCAGCTATTATTTGCTGATGTGCAAGAGCGTACCCTGTAACACGCGNA	360
OY	361	GCGTTGCGCGCGGCGCTGCGACCTACTGTGCAATTAACCGTACATACCAGCAATACG	420
DB	361	GCGTTGCGCGCGGCGCTGCGACCTACTGTGCAATTAACCGTACATACCAGCAATACG	420
OY	421	TCGAGTCCCGAATCCGTTATTTAATGCTCTTAAATAATGGCGTTTGCCAACTGGATTAACGG	480
DB	421	TCGAGTCCCGAATCCGTTATTTAATGCTCTTAAATAATGGCGTTTGCCAACTGGATTAACGG	480
OY	481	AACGTGACTGACGCGATCTTCAGCAGGCGCAGAGGGTCAATTGCTGACTTTACCGGGCAT	540
DB	481	AACGTGACTGACGCGATCTTCAGCAGGCGCAGAGGGTCAATTGCTGACTTTACCGGGCAT	540
OY	541	CGGCAAAAGGGGTTTCGGGAATGGAAGGGGCTTAATTTCCGCATCAACTGATGTC	600
DB	541	CGGCAAAAGGGGTTTCGGGAATGGAAGGGGCTTAATTTCCGCATCAACTGATGTC	600
OY	601	CTTAAACGTGAGAAACAGAGCAAGAAAGCTTTCATTAACGACAGGCAATTAACATCGAAGTTC	660
DB	601	CTTAAACGTGAGAAACAGAGCAAGAAAGCTTTCATTAACGACAGGCAATTAACATCGAAGTTC	660
OY	661	AAGGTGAGCGCGGACATGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG	720
DB	661	AAGGTGAGCGCGGACATGTCTCATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG	720
OY	721	GAGATATTTCTCTGCAACAAAGCAACAGGAAATGCCGAGCGGGGTGGGAAGGATCAAC	780
DB	721	GAGATATTTCTCTGCAACAAAGCAACAGGAAATGCCGAGCGGGGTGGGAAGGATCAAC	780
OY	781	GATTACACACAGTGGAAACACCTGTGTAGTTTGGCTAATACGCGCAATTTTATTTGCTACAA	840
DB	781	GATTACACACAGTGGAAACACCTGTGTAGTTTGGCTAATACGCGCAATTTTATTTGCTACAA	840
OY	841	CGCAGCGCAGAGGTGGCCGCGACGCCGCCACCCGTTATTGGATTGATCATGTGCACG	900
DB	841	CGCAGCGCAGAGGTGGCCGCGACGCCGCCACCCGTTATTGGATTGATCATGTGCACG	900
OY	901	TTGAGCGCCCAATCCACCGCAAAACAGCGATGTGTGGATGACATTAACCACTTCACTACTG	960
DB	901	TTGAGCGCCCAATCCACCGCAAAACAGCGATGTGTGGATGACATTAACCACTTCACTACTG	960

Qy	961	TTTTATGCGGACACGATTAATCTGGCAATCTCGGCGGCGACATGAGACTCAACTGG	1020		
Db	961	TTTATTTGCGGACACGATTAATCTGGCAATCTCGGCGGCGACATGAGACTCAACTGG	1020		
Qy	1021	ACGCTTCCCGGTCACGCGGATTAACACGCGCCAGAGTGATGTAAGTGTGTTGAACGCTGG	1080		
Db	1021	ACGCTTCCCGGTCACGCGGATTAACACGCGCCAGAGTGATGTAAGTGTGTTGAACGCTGG	1080		
Qy	1081	CGTGGGCTTAAGCGTAACAGCCAGTGGATTAGGTTTCGCTGGTCTTCAGACTTTTACAG	1140		
Db	1081	CGTGGGCTTAAGCGTAACAGCCAGTGGATTAGGTTTCGCTGGTCTTCAGACTTTTACAG	1140		
Qy	1141	CAGATGCGGTAAAGCGCGCTGTCAATTAATACGCGCGCCGAGAGAGTGAACCTGAC	1200		
Db	1141	CAGATGCGGTAAAGCGCGCTGTCAATTAATACGCGCGCCGAGAGAGTGAACCTGAC	1200		
Qy	1201	CTGGCAGAGATGTGAAGAGCGCAATGCGCAGGCGATGTGTTGTTGGCAGGTTTACGCA	1260		
Db	1201	CTGGCAGAGATGTGAAGAGCGCAATGCGCAGGCGATGTGTTGTTGGCAGGTTTACGCA	1260		
Qy	1261	ATCGTGAATGAAGACCGCATACCGCGGTGGCAGTTTGAGATCTCATACCATCACCATCAC	1320		
Db	1261	ATCGTGAATGAAGACCGCATACCGCGGTGGCAGTTTGAGATCTCATACCATCACCATCAC	1320		
Qy	1321	TAA 1323			
Db	1321	TAA 1323			
RESULT 2	ARI27818	1323 bp	DNA	Linear	PAT 16-MAY-2001
LOCUS	ARI27818	Sequence 1 from patent US 6183740.			
DEFINITION	ARI27818				
ACCESSION	ARI27818				
VERSION	ARI27818.1	GI:14115480			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1323)				
AUTHORS	Short, J.M. and Kretz, K.A.				
TITLE	Recombinant bacterial phylases and uses thereof				
JOURNAL	Patent: US 6183740-A 1 06-FEB-2001;				
FEATURES	Location/Qualifiers				
source	1..1323				
BASE COUNT	323 a 353 c 357 g 289 t			1 others	
ORIGIN					
Query Match	99.9%; Score 1322; DB 6; Length 1323;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 1323;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGAAGCGATCTTAATCCATTTTATATCTTGTGATTCGGTTAAACCCGCAATCTGCA	60		
Db	1	ATGAAGCGATCTTAATCCATTTTATATCTTGTGATTCGGTTAAACCCGCAATCTGCA	60		
Qy	61	TTTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGGTGATGTGCAGTGCATAGT	120		
Db	61	TTTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGGTGATGTGCAGTGCATAGT	120		
Qy	121	GTGCGTCTCCAAACCAAGGCGACGCACTGATGCAGATGTCAACCCAGACGATGGCCA	180		
Db	121	GTGCGTCTCCAAACCAAGGCGACGCACTGATGCAGATGTCAACCCAGACGATGGCCA	180		
Qy	181	ACGTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAACTGCTATCTC	240		
Db	181	ACGTGGCCGGTAAACTGGGTTGGCTGACACCGCGAGGTGGTGAAGTAACTGCTATCTC	240		
Qy	241	GGAATTAACCAACCCGAGCGTGTGGTGAAGCGGATGCTGGCGGAAAGAGCGTGGCCG	300		
Db	241	GGAATTAACCAACCCGAGCGTGTGGTGAAGCGGATGCTGGCGGAAAGAGCGTGGCCG	300		

Db	721	GAGATATTTCCTCGCAACAAGCAGCAGGATGCGGAGCCGGGGTGGGAGAGATCACC	780
QY	781	GATTACACCCAGTGGAAACACCTTGCTAAGTTTGCAATAACCGCAATTTATTTGCTACAA	840
Db	781	GATTACACACCGTGGAAACACCTTCTAAGTTTGCAATAACCGCAATTTATTTATTTCTACAA	840
QY	841	CGCAGCGCAGAGGTGGCCCGCAGCGGCCACCCCGTATTGATTTGATATGATGGCAGCG	900
Db	841	CGCAGCGCAGAGGTGGCCCGCAGCGGCCACCCCGTATTGATTTGATATGATGGCAGCG	900
QY	901	TTGACGCCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG	960
Db	901	TTGACGCCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG	960
QY	961	TTTATTGCGGACACGATACTAATCTGGCAAAATCTGGCGCGCGCAGTGGAGCTCAACTG	1020
Db	961	TTTATTGCGGACACGATACTAATCTGGCAAAATCTGGCGCGCGCAGTGGAGCTCAACTG	1020
QY	1021	ACGCTTCCCGGTACGCGGATTAACAGCCCGCCAGAGTGTGAACGTGTGGAACGCTGG	1080
Db	1021	ACGCTTCCCGGTACGCGGATTAACAGCCCGCCAGAGTGTGTGAACGTGTGGAACGCTGG	1080
QY	1081	CGTGGCGTAAAGCGTAACAGCCAGTTCAGTTTCGCTGCTGCTTCACACTTTTACAG	1140
Db	1081	CGTGGCGTAAAGCGCGCTGTCAATTAATACGCGCCCGGAGAGAGTGAACCTGACC	1140
QY	1141	CAGATGGCGTAAAGCGCGCTGTCAATTAATACGCGCCCGGAGAGAGTGAACCTGACC	1200
Db	1141	CAGATGGCGTAAAGCGCGCTGTCAATTAATACGCGCCCGGAGAGAGTGAACCTGACC	1200
QY	1201	CTGGCAGAGTGTGAAGAGCGAAATGCGCAGGGCATGTGTTCTGGCGAGTTTACGCA	1260
Db	1201	CTGGCAGAGTGTGAAGAGCGAAATGCGCAGGGCATGTGTTCTGGCGAGTTTACGCA	1260
QY	1261	ATCGTGAATGAAGACCATATCCGCGCTGCAGTTTGAGATTCATCAACATCACCATCAC	1320
Db	1261	ATCGTGAATGAAGACCATATCCGCGCTGCAGTTTGAGATTCATCAACATCACCATCAC	1320
QY	1321	TAA 1323	
Db	1321	TAA 1323	
RESULT 4			
AX052713			
LOCUS	AX052713	1323 bp	DNA
DEFINITION	Sequence 1 from Patent WO0071728.	linear	PAT 12-JAN-2001
ACCESSION	AX052713		
VERSION	AX052713.1	GI:12226903	
KEYWORDS			
SOURCE			
ORGANISM	Escherichia coli.		
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE	1 (bases 1 to 1323)		
AUTHORS	Short, J.M. and Kretz, K.A.		
TITLE	Recombinant bacterial phytases and uses thereof		
JOURNAL	Patent: WO 0071728-A 1 30-NOV-2000; Diversa Corporation (US)		
FEATURES			
source	1..1323	Location/Qualifiers	
	/organism="Escherichia coli"		
	/db_xref="taxon:562"		
	1..1323		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=1		
	/protein_id="CAC21810.1"		
	/db_xref="GI:12226904"		
	/translation="MKAILIPFLSLPLTPQSAFASPEPKIESVVIYSRHEVRAP		
	TKATQLMODVTPDAMPTWPKLGLWTRPGEGLNATLGHYQRRLVADGLAKKCPQSS		
	GOVAIIADVETRRKTEAFPAAGIAPCAITVHQATQSSDPFLNPLKTVGVCIDNA		

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Short, J. M., Kretz, R. A. and O'Donoghue, E.	Dietary aids and methods of use therefor	Patent: WO 0189317-A 1 29-NOV-2001; DIVERSA CORPORATION (US)

FEATURES	Location/Qualifiers
source	1. .1323

```

/organism="Escherichia coli"
/db_xref="taxon:562"
1. .1323
CDS

```

```

/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD20553.1"
/db_xref="GI:18129103"
/translation="MKAILIPFLSLPIPLIPQSAFAQSEPEELKESVIVSRHGVRAP
TKAQLMODVTPDPMAPVVKLGMVLPFGGELIAYLGHYORRLVADGLLAKKCPQS
GOVVALIADVDERPRTKEAFAPAGLAPPCATVTHQAQTSPPDLFENLPKGYCQIDNNA
NMTDAIILSRAGGSTADFTGHRQIAFRELERYLANPQSNILKREKQDEQSCSLTQALPS
ELKASADVNSLITGAVNSLASMLTEILFLEQLAQOGMEPEWGRITRSHQWNTLTLSNAPV
YLQRTPEVARSRAQTPLLDLIMALTPEHPQOKQGVTLPLPSVFIAGHPTNLNLKLGTC
ALLENMTLPGQPDNTPPGGELVFERMRRLSPNSQOIVSLVFOFLQOMRDKTPLSLMTV
PPGKGVKTLACGERRNAQGMCSLAGFPQIYNEARIPRCSLSRHHHHH"
BASE COUNT      333 a      353 c      357 g      289 t      1 others
ORIGIN

```

Query Match	99.9%	Score 1322;	DB 6;	Length 1323;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1323;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

```

|||||
Db      .1081  CGTGGGCTAAAGCGATTAACAGCCAGTGATTCAGGTTTGTGCTGCTTCCAGACTTTACAG 1140
Oy      1141  CAGATGGCTGATTAAGCCGCTGTCATTAATACGCCGCCGAGAGGTAAATGACC 1200
Db      1141  CAGATGGCTGATTAAGCCGCTGTCATTAATACGCCGCCGAGAGGTAAATGACC 1200
Oy      1201  CTGGCAGGATGTGAAGAGCGAATGCGCAGGCGCATGTGCTGGCGAGTTTACGCA 1260
Db      1201  CTGGCAGGATGTGAAGAGCGAATGCGCAGGCGCATGTGCTGGCGAGTTTACGCA 1260
Oy      1261  ATCGTGATGAAGACGCGATACCGCGCTGCGAGTTTGAGATCTCATCACCATCACCATAC 1320
Db      1261  ATCGTGATGAAGACGCGATACCGCGCTGCGAGTTTGAGATCTCATCACCATCACCATAC 1320
Oy      1321  TAA 1323
Db      1321  TAA 1323
Db      1321  TAA 1323

RESULT 6
AX356566 1323 bp DNA linear PAT 06-FEB-2002
LOCUS     Sequence 1 from Patent WO0190333.
ACCESSION AX356566
VERSION   AX356566.1 GI:18621053
KEYWORDS
SOURCE    Escherichia coli.
ORGANISM  Escherichia coli.
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
REFERENCE
  1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrell, J.B. and
    O'Donoghue, E.
    Recombinant bacterial phylases and uses thereof
    Patent: WO 0190333-A 1 29-NOV-2001;
    DIVERSA CORPORATION (US)
FEATURES
  source
    1..1323
    /location="Qualifiers
    /organism="Escherichia coli"
    /db_xref="taxon:562"
    1..1323
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAD23025.1"
    /db_xref="GI:18621054"
    /translation="MKAILIPFLSLIPFQSAFQSEPELKIYESVIVSRHGVRA
    TKATIMODVTPDAMPVPRKILWLTIPRGELIAYLGHYQKRLVADLLAKKCPQS
    GQVATIDVDERTRKTEGAFNAGLAPDCATVHTQADSSPDLPNPLKTCVQOLDNA
    NVTDAIISRAGSIADEFTGHRQTAFLRELELVNFPQSMCLKREKODSCSLTQALPS
    ELKVSADNVSLTGAVSLASMLTEIFLLQOAGMEPGMGRTSDSHQWTLSTLNAQF
    YLQRTPEVARSRATPLDLIMALTPLPPOKQAGVLTPLTSYLEIAGHDNINLIG
    ALEMTLPGQDNTPPGGLVFERMRRLSDNSOMIOVSLFQTLQQRDKTPIPLSNT
    PPGREYKTLIACPERNMGKSLAGFTQIVNRIAPRACSLSHHHHHH"
BASE COUNT 323 a 353 c 357 g 289 t 1 others
ORIGIN
Query Match 99.9%; Score 1322; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      181  ACCTGGCCGTAATAACTGGTGTGCTGACACCGCGNGTGTGAGCTAATGCGCTATCTC 240
Db      181  ACCTGGCCGTAATAACTGGTGTGCTGACACCGCGNGTGTGAGCTAATGCGCTATCTC 240
Oy      241  GGACATTAACAAAGCCAGCGCTGTGTAGCCGACGATTTGCTGGCAAAAAGGCTGCCG 300
Db      241  GGACATTAACAAAGCCAGCGCTGTGTAGCCGACGATTTGCTGGCAAAAAGGCTGCCG 300
Oy      301  CAGTGTGTCAGGTGCGCATTAATGCTGATGCGAGAGCGCATACCGTAAACAGCGCA 360
Db      301  CAGTGTGTCAGGTGCGCATTAATGCTGATGCGAGAGCGCATACCGTAAACAGCGCA 360
Oy      361  GCGTTGCGCGCGCGGCTGGACACTGACATGTGCAATACCGTACATACCGAGCAGATACG 420
Db      361  GCGTTGCGCGCGCGGCTGGACACTGACATGTGCAATACCGTACATACCGAGCAGATACG 420
Oy      421  TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGCGCTTGGCAACTGATTAACGG 480
Db      421  TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGCGCTTGGCAACTGATTAACGG 480
Oy      481  AACGTGACGACGCGATCCTCAGAGGAGGAGGCTCAATTCGTGACTTTACCGGGCAT 540
Db      481  AACGTGACGACGCGATCCTCAGAGGAGGAGGCTCAATTCGTGACTTTACCGGGCAT 540
Oy      541  CGGCAAAAGCGGCTTTCGGGAACGTGACGCGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Db      541  CGGCAAAAGCGGCTTTCGGGAACGTGACGCGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Oy      601  CTTAAACGTGAGAAACAGAGCAAGCTGTTCATTAAACGCGCATTAACCGGAACTC 660
Db      601  CTTAAACGTGAGAAACAGAGCAAGCTGTTCATTAAACGCGCATTAACCGGAACTC 660
Oy      661  AAGGTGACGCGCGGACAAATGTCATTAAACCGGTGCGTAAGCCTGCAATATGCTGACG 720
Db      661  AAGGTGACGCGCGGACAAATGTCATTAAACCGGTGCGTAAGCCTGCAATATGCTGACG 720
Oy      721  GAGATATTTCTCTGCAACACAGCAGGATTCGGAGCCGGGTGGGGAAGATCAC 780
Db      721  GAGATATTTCTCTGCAACACAGCAGGATTCGGAGCCGGGTGGGGAAGATCAC 780
Oy      781  GATTCACACCGTGGAAACACCTTGTGATTAACGCGCATTTATTTGCTACAA 840
Db      781  GATTCACACCGTGGAAACACCTTGTGATTAACGCGCATTTATTTGCTACAA 840
Oy      841  CGCAGCGCAGAGTTCGCCGCGACCGCGCACCCGCTTAATTTGATTTGATCAATGCGACG 900
Db      841  CGCAGCGCAGAGTTCGCCGCGACCGCGCACCCGCTTAATTTGATTTGATCAATGCGACG 900
Oy      901  TTGACGCCCCATCCACCGCAAAACAGCGCTATGTGTGACATTAACCACTTCACTACTG 960
Db      901  TTGACGCCCCATCCACCGCAAAACAGCGCTATGTGTGACATTAACCACTTCACTACTG 960
Oy      961  TTTATATGCGGACAGATTAATCTGCAATTCGCGCGGCGGCACTGGAGCTCAACTGG 1020
Db      961  TTTATATGCGGACAGATTAATCTGCAATTCGCGCGGCGGCACTGGAGCTCAACTGG 1020
Oy      1021  ACGCTTCCCGGTCACCGGATTAACAGCGCGCGCAGGTGTGAATGCTGTTTGAACGTG 1080
Db      1021  ACGCTTCCCGGTCACCGGATTAACAGCGCGCGCAGGTGTGAATGCTGTTTGAACGTG 1080
Oy      1081  CGTGGCTTAAGCGATTAACAGCGCATGTGATTCAGTTTGTGCTTCCAGACTTTACG 1140
Db      1081  CGTGGCTTAAGCGATTAACAGCGCATGTGATTCAGTTTGTGCTTCCAGACTTTACG 1140
Oy      1141  CAGATGCGTGATTAAGCCGCTGTCATTAATACGCCGCCGAGAGAGTGAATAATGACC 1200
Db      1141  CAGATGCGTGATTAAGCCGCTGTCATTAATACGCCGCCGAGAGAGTGAATAATGACC 1200
Oy      1201  CTGGCAGGATGTGAAGAGCGAATGCGCAGGCGCATGTGCTGGCGAGTTTACGCA 1260
Db      1201  CTGGCAGGATGTGAAGAGCGAATGCGCAGGCGCATGTGCTGGCGAGTTTACGCA 1260

```

QY 1261 ATCGTGATGAGACGACCATACCGGCGTGCAGTTTGAGATCTCATACCATCACCACATC 1320
 Db 1261 ATCGTGATGAGACGACCATACCGGCGTGCAGTTTGAGATCTCATACCATCACCACATC 1320
 QY 1321 TAA 1323
 Db 1321 TAA 1323
 RESULT 7
 ECOAPPAA 1901 bp DNA linear BCT 04-APR-2002
 LOCUS Escherichia coli periplasmic phosphoanhydride phosphohydrolase
 DEFINITION (appa) gene, complete cds.
 ACCESSION M58708
 VERSION M58708.1 GI:145283
 KEYWORDS Escherichia coli.
 SOURCE Escherichia coli.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia
 REFERENCE 1 (bases 1 to 1901)
 AUTHORS Dassa, J., Marck, C. and Boquet, P. L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appa
 reveals significant homology between pH 2.5 acid phosphatase and
 glucose-1-phosphatase
 J. Bacteriol. 172 (9), 5497-5500 (1990)
 JOURNAL MEDLINE 90368616
 PUBMED 2168385
 REFERENCE 2 (sites)
 AUTHORS Ostanin, K., Harms, E. H., Stevis, P. E., Kuciel, R., Zhou, M. M. and Van
 Etten, R. L.
 TITLE Overexpression, site-directed mutagenesis, and mechanism of
 Escherichia coli acid phosphatase
 J. Biol. Chem. 267 (32), 22830-22836 (1992)
 JOURNAL MEDLINE 93054596
 PUBMED 1429631
 REFERENCE 3 (sites)
 AUTHORS Ostanin, K. and Van Etten, R. L.
 TITLE App304 of Escherichia coli acid phosphatase is involved in leaving
 group protonation
 J. Biol. Chem. 268 (28), 20778-20784 (1993)
 JOURNAL MEDLINE 94012612
 PUBMED 8407904
 FEATURES
 source location/Qualifiers
 1..1901 /organism="Escherichia coli"
 /db_xref="taxon:562"
 3..1522 /gene="appa"
 3..8 /gene="appa"
 16..108 /gene="appa"
 /note="ORF"
 /codon_start=1
 /transl_table=11
 /product="unknown"
 /protein_id="AAI2085.1"
 /db_xref="GI:145284"
 /translation="MYLLWFVGIILMCSLSITLVLVWLPRLKS"
 118..149 /gene="appa"
 133..138 /gene="appa"
 156..161 /gene="appa"
 172..177 /gene="appa"
 188..1486 /gene="appa"
 /gene="appa"
 /standard_name="acid phosphatase"
 /codon_start=1
 stem_loop
 -35_signal
 -10_signal
 RBS
 CDS

/transl_table=11
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AAI2085.1"
 /db_xref="GI:145283"
 /translation="MAKILIPFLSLIPLPQSAQSEPELKESVIVSRHGVAP
 TKATQIMQDVTYPMWVYGLWLTFRGGLIYLVGHYQORLVADGLAKGCPQS
 GOVALIADVDERTKTEGEAFAGLAPCAITVHQADTSSPDLFPLKGVCGLDNA
 MYDAILISRAAGSITADFTGHRQTAFLRELYINFPQSNICIKREKQESCTIQA
 ELKYSADNVSLITGAVSLASMLTEIFLLQAOAGMPEWGRITDSHNNLTLSLNAQF
 VLIQRTPEVARSRAVPLDLITKLTALTPHPQKQAVGYVLPVSLFVIGHDNTNANLGG
 ALIELNTWLPQGPDMTPPGGELVFERMRRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNT
 PGCEVKTLIAGCEERNAQMCISLQIYVIEARI PACSL"
 188..253 /gene="appa"
 254..1483 /gene="appa"
 /product="periplasmic phosphoanhydride phosphohydrolase"
 299..301 /gene="appa"
 /standard_name="H303a mutant"
 /note="created by site directed mutagenesis"
 /citation=[2]
 299..301 /gene="appa"
 /standard_name="R16a mutant"
 /note="created by site directed mutagenesis"
 /citation=[2]
 299..301 /gene="appa"
 /standard_name="R20a mutant"
 /note="created by site directed mutagenesis"
 /citation=[2]
 299..301 /gene="appa"
 /standard_name="H17N mutant"
 /note="created by site directed mutagenesis"
 /citation=[2]
 302..304 /gene="appa"
 /standard_name="D304E mutant"
 /note="created by site directed mutagenesis"
 /citation=[2]
 1163..1165 /gene="appa"
 /standard_name="D304A mutant"
 /note="created by site directed mutagenesis"
 /citation=[3]
 /phenotype="Asp 304 changed to Ala"
 /replace="gct"
 1163..1165 /gene="appa"
 /standard_name="D304E mutant"
 /note="created by site directed mutagenesis"
 /citation=[3]
 /phenotype="Asp 304 changed to Glu"
 /replace="gag"
 1497..1522 /gene="appa"
 /note="rho-independent terminator"
 terminator
 BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN
 Query Match 97.3%; Score 1287.8; DB 1; Length 1901;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAAGGAGATCTAATCCATTTTATCTTCGTGATCGGTTAACCCGCAATTCGA 60
 Db 188 ATGAAGGAGATCTAATCCATTTTATCTTCGTGATCGGTTAACCCGCAATTCGA 247
 QY 61 TTCGCTGAGAGTGAAGCCGAGCTGAAGGTGAGTGTGATGTCAGTCGTCATGCT 120
 Db 248 TTCGCTGAGATGAGCCGAGCTGAAGGTGAGTGTGATGTCAGTCGTCATGCT 307

QY 121 GTGCGTGTCCAAAGGCGACGCACTGATGCAAGATGTCACCCAGACGCAATGCGCA 180
 DB 308 GTGCGTGTCCAAAGGCGACGCACTGATGCAAGATGTCACCCAGACGCAATGCGCA 367
 QY 181 ACCGTGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGGTGAAGTAAATGCTATCTC 240
 DB 368 ACCGTGCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGGTGAAGTAAATGCTATCTC 427
 QY 241 GGACATTTACCAAGCGCAGCGTGTGGTAGCGGATTCCTGGCGGAAAAAGGGCTGCCG 300
 DB 428 GGACATTTACCAAGCGCAGCGTGTGGTAGCGGATTCCTGGCGGAAAAAGGGCTGCCG 487
 QY 301 CAGCTGTGTGAGTGTGCGGATTAATGCTGATGTCAGACGAGCGTAAACCGGCGAA 360
 DB 488 CAGCTGTGTGAGTGTGCGGATTAATGCTGATGTCAGACGAGCGTAAACCGGCGAA 547
 QY 361 GCGTTGCGCGCGCGGCGTGGCAGCTGTGCAATTAACCGTAAATCCAGGCAATAG 420
 DB 548 GCGTTGCGCGCGCGGCGTGGCAGCTGTGCAATTAACCGTAAATCCAGGCAATAG 607
 QY 421 TCACATCCGATCCGTTATTTATCTCTAAACCTGGGTTGGCAACTGGATAACGCG 480
 DB 608 TCACATCCGATCCGTTATTTATCTCTAAACCTGGGTTGGCAACTGGATAACGCG 667
 QY 481 AACGTGACTGACGCGATCTCTAGCAGGCGAGGCGTCAATTGCTGACTTTACCGGCGAT 540
 DB 668 AACGTGACTGACGCGATCTCTAGCAGGCGAGGCGTCAATTGCTGACTTTACCGGCGAT 727
 QY 541 CGGCAAAAGCGGTTTCGGCACTGCAAGCGGCTTAATTTTCCGCAATCAACTTGTGC 600
 DB 728 CGGCAAAAGCGGTTTCGGCACTGCAAGCGGCTTAATTTTCCGCAATCAACTTGTGC 787
 QY 601 CTTAAACGTGAGAAACAGAGCAGAACTGTCTATTAAACGACGCAATTAACCATCGAATCTC 660
 DB 788 CTTAAACGTGAGAAACAGAGCAGAACTGTCTATTAAACGACGCAATTAACCATCGAATCTC 847
 QY 661 AAGGTGAGCGCGGACAACTGTCTCAATTAACCGGCTGGGTAAGCCTGCAATCAATGCTGACG 720
 DB 848 AAGGTGAGCGCGGACAACTGTCTCAATTAACCGGCTGGGTAAGCCTGCAATCAATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACACAGCAGGAAATCCGCGGACCGGCGGTGGGGAAGATACAC 780
 DB 908 GAGATATTTCTCTGCAACACAGCAGGAAATCCGCGGACCGGCGGTGGGGAAGATACAC 967
 QY 781 GATTTCACACAGTGGAAACCTTGTGTAAGTTGCAATTAACGCGCAATTTTATTGCTACAA 840
 DB 968 GATTTCACACAGTGGAAACCTTGTGTAAGTTGCAATTAACGCGCAATTTTATTGCTACAA 1027
 QY 841 CGCAGCGCAGAGGTTGCGCGACGCGCGCACCCCGTTATTTGATTTGATGCGACGCG 900
 DB 1028 CGCAGCGCAGAGGTTGCGCGACGCGCGCACCCCGTTATTTGATTTGATGCGACGCG 1087
 QY 901 TTGAGAGCGCCATCCACCGCAAAAAAAGAGCGTATGCTGTGACATTAACCCACTTCACTACG 960
 DB 1088 TTGAGAGCGCCATCCACCGCAAAAAAAGAGCGTATGCTGTGACATTAACCCACTTCACTACG 1147
 QY 961 TTTATTTGCGGACAGATCTAATCTGCGCAATCTGCGGCGGCACTGGAGCTCAACTGCG 1020
 DB 1148 TTTATTTGCGGACAGATCTAATCTGCGCAATCTGCGGCGGCACTGGAGCTCAACTGCG 1207
 QY 1021 ACGCTTCCCGGTACGCGGATTAACAGCGCGCGAGGTGTGAACGTGTTTGAACGCTGG 1080
 DB 1208 ACGCTTCCCGGTACGCGGATTAACAGCGCGCGAGGTGTGAACGTGTTTGAACGCTGG 1267
 QY 1081 CGTGGGCTAAAGGATTAACAGCGCAGTGTGATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1140
 DB 1268 CGTGGGCTAAAGGATTAACAGCGCAGTGTGATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1327
 QY 1141 CAGATGCGTGAATAACGCGCGTCTATTAAATAGCGCGCGGAGAGGTGAACCTGACC 1200
 DB 1328 CAGATGCGTGAATAACGCGCGTCTATTAAATAGCGCGCGGAGAGGTGAACCTGACC 1387

QY 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTCTGTTGGCAGGTTTACGCA 1260
 DB 1388 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTCTGTTGGCAGGTTTACGCA 1447
 QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTCCACTTTGAGAT 1300
 DB 1448 ATCGTGAATGAAGCAGCATACCGGCGTCCACTTTGAGAT 1487
 RESULT 8
 AX356572 1901 bp DNA linear PAT 06-FEB-2002
 LOCUS AX356572
 DEFINITION Sequence 7 from Patent WO0190333.
 ACCESSION AX356572
 VERSION AX356572.1 GI:18621059
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.
 Recombinant bacterial phylases and uses thereof
 Patent: WO 0190333-A 7 29-NOV-2001;
 DIVERSA CORPORATION (US)
 FEATURES
 source
 1. 1901
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN
 Query Match 97.3%; Score 1287.8; DB 6; Length 1901;
 Best local similarity 99.5%; Pred. No. 0;
 Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTGATTCGTTAAACCGCAATCTGCA 60
 DB 188 ATGAAGCGATCTTAATCCCAATTTTATCTCTGATTCGTTAAACCGCAATCTGCA 247
 QY 61 TTGCGTCAGAGTGAAGCGCGGAGCTGAAGCTGGAAGTGTGTGATGTCTGCTCATGT 120
 DB 248 TTGCGTCAGAGTGAAGCGCGGAGCTGAAGCTGGAAGTGTGTGATGTCTGCTCATGT 307
 QY 121 GTGCGTGTCCAAAGGCGACGCACTGATGCAAGATGTCACCCAGACGCAATGCGCA 180
 DB 308 GTGCGTGTCCAAAGGCGACGCACTGATGCAAGATGTCACCCAGACGCAATGCGCA 367
 QY 181 ACCGTGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGTGAGCTAATTCGCTATCTC 240
 DB 368 ACCGTGCGGTAAACCTGGGTGGCTGACACCGCGNGGTGTGAGCTAATTCGCTATCTC 427
 QY 241 GGACATTTACCAAGCGCAGCGTGTGTGACCGGACGATTCCTGGCGGAAAAAGGGCTGCCG 300
 DB 428 GGACATTTACCAAGCGCAGCGTGTGTGACCGGACGATTCCTGGCGGAAAAAGGGCTGCCG 487
 QY 301 CAGCTGTGTGAGTGTGCGGATTAATGCTGATGTCAGACGAGCGTAAACCGGCGAA 360
 DB 488 CAGCTGTGTGAGTGTGCGGATTAATGCTGATGTCAGACGAGCGTAAACCGGCGAA 547
 QY 361 GCGTTGCGCGCGGCGTGGCAGCTGTGCAATTAACCGTAAATCCAGGCAATAG 420
 DB 548 GCGTTGCGCGCGGCGTGGCAGCTGTGCAATTAACCGTAAATCCAGGCAATAG 607
 QY 421 TCACATCCGATCCGTTATTTATCTCTAAACCTGGGTTGGCAACTGGATAACGCG 480
 DB 608 TCACATCCGATCCGTTATTTATCTCTAAACCTGGGTTGGCAACTGGATAACGCG 667
 QY 481 AACGTGACTGACGCGATCTCTAGCAGGCGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 540
 DB 668 AACGTGACTGACGCGATCTCTAGCAGGCGAGGAGGCTCAATTGCTGACTTTACCGGCGAT 727

```

QY 541 CGGCAACGGCGTTTCGCAACTGGAACGGGTCCTAATTTTCCGCAATCAAACTGTGC 600
    |||||||
Db 728 CGGCAACGGCGTTTCGCAACTGGAACGGGTCCTAATTTTCCGCAATCAAACTGTGC 787
QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGCATTACCAATCGAACTC 660
    |||||||
Db 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGCATTACCAATCGAACTC 847
QY 661 AAGGTGAGCGCCGCAATGTCATTAACCGGTGCGGTAAAGCCGCAATCAATCTGACG 720
    |||||||
Db 848 AAGGTGAGCGCCGCAATGTCATTAACCGGTGCGGTAAAGCCGCAATCAATCTGACG 907
QY 721 GAGATATTTCTCTGCAACAAGACAGAGGATGCGGAGCGGCGGTGGGAGATGATACC 780
    |||||||
Db 908 GAGATATTTCTCTGCAACAAGACAGAGGATGCGGAGCGGCGGTGGGAGATGATACC 967
QY 781 GATTACACACAGTGGAAACACCTTGCTAAGTTGCATAACCGCAATTTTATTTGCTACAA 840
    |||||||
Db 968 GATTACACACAGTGGAAACACCTTGCTAAGTTGCATAACCGCAATTTTATTTGCTACAA 1027
QY 841 CGGACGCGCAGAGTTGCGCGCAGCGCGCACCCCGTTATTTGATTTGATCATGCGACGC 900
    |||||||
Db 1028 CGGACGCGCAGAGTTGCGCGCAGCGCGCACCCCGTTATTTGATTTGATCATGCGACGC 1087
QY 901 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 960
    |||||||
Db 1088 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 1147
QY 961 TTTATTCGCGACACAGTACTTAATCTGGCAAAATCTGCGCGCGCAGCTGAGACTCACTG 1020
    |||||||
Db 1148 TTTATTCGCGACACAGTACTTAATCTGGCAAAATCTGCGCGCGCAGCTGAGACTCACTG 1207
QY 1021 ACGGTTCCGGGTGACGCGGATTAACAGCGCGCAGGTGATGACGTGTTTGAACGCTGC 1080
    |||||||
Db 1208 ACGGTTCCGGGTGACGCGGATTAACAGCGCGCAGGTGATGACGTGTTTGAACGCTGC 1267
QY 1081 CGTGCGGTAAGCGATTAACAGCGAGTGTTCAGGTTTCGCTGCTCCAGACTTTACAG 1140
    |||||||
Db 1268 CGTGCGGTAAGCGATTAACAGCGAGTGTTCAGGTTTCGCTGCTCCAGACTTTACAG 1327
QY 1141 CAGATGCGGTATAAAGCGCGCTGTCAATTAATACGCCGCCGAGAGGTGAATGACACC 1200
    |||||||
Db 1328 CAGATGCGGTATAAAGCGCGCTGTCAATTAATACGCCGCCGAGAGGTGAATGACACC 1387
QY 1201 CTTGCGAGAGTGTGAAGCGCAAAATGGCAGGCGCATGCTGCTGCGAGGTTTACCAA 1260
    |||||||
Db 1388 CTTGCGAGAGTGTGAAGCGCAAAATGGCAGGCGCATGCTGCTGCGAGGTTTACCAA 1447
QY 1261 ATCGTGAATGAAGACGACATACCGCGCTGCAGTTTGAGAT 1300
    |||||||
Db 1448 ATCGTGAATGAAGACGACATACCGCGCTGCAGTTTGAT 1487

```

```

RESULT 9
AE000200
LOCUS Escherichia coli K12 MG1655 section 90 of 400 of the complete
DEFINITION genome.
ACCESSION AE000200 U00096
VERSION AE000200.1 GI:2367111
KEYWORDS
SOURCE
ORGANISM Escherichia coli K12.
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
REFERENCE 1 (bases 1 to 10784)
          Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
          Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
          Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J.,
          Mau, B. and Shao, Y.
          The complete genome sequence of Escherichia coli K-12
          Science 277 (5331), 1453-1474 (1997)
TITLE JOURNAL
JOURNAL MEDLINE
97426617

```

```

PUBMED 9278503
REFERENCE 2 (bases 1 to 10784)
AUTHORS Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
REFERENCE 3 (bases 1 to 10784)
AUTHORS Blattner, F.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (bases 1 to 10784)
Plunkett, G. III.
Direct Submission
Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
On Sep 9, 1997 this sequence version replaced gi:1787214.
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCHGR). The entire sequence was independently
determined from E. coli K12 strain MG1655. Predicted open reading
frames were determined using Genemark software, kindly supplied by
Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
30332 (e-mail: markborodov.gatech.edu). Open reading frames that
have been correlated with genetic loci are being annotated with CG
site Nos., unique ID nos. for the genes in the E. coli Genetic
Stock Center (CGSC) database at Yale University, kindly supplied by
Mary Berlyn. A public version of the database is accessible
(http://cgsc.biology.yale.edu). Annotation of the genome is an
ongoing task whose goal is to make the genome sequence more useful
by correlating it with other data. Comments to the authors are
appreciated. Updated information will be available at the E. coli
Genome Project's World Wide Web site
(http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
its annotations are periodically updated; this is version M54. No
sequence changes. Annotation updates: updated gene identifications
and products; all new functional assignments courtesy of Monica
Riley; added promoters, protein binding sites, and repeated
sequences described in reference 1. The unique numeric identifiers
beginning with a lowercase 'b' assigned to each gene (protein- or
RNA-encoding) are now designated as gene synonyms instead of
labels. This should allow them to be searched for in Entrez as gene
names.
FEATURES
source location/Qualifiers
1..10784
/organism="Escherichia coli K12"
/strain="K12"
/sub_strain="MG1655"
/db_xref="taxon:83333"
49..84
/note="REP (repetitive extragenic palindromic) element;
contains 1 REP sequence"
123..1421
/gene="appa"
/note="b0980"
123..1421
/gene="appa"
/EC_number="3.1.3.2"
/EC_number="3.1.3.26"
/function="enzyme; Central intermediary metabolism: Pool,
multipurpose conversions"
/note="0432; 100 pct identical to PPA_ECOLI SW: P07102"
/codon_start=1
/transl_table=11
/product="phosphoanhydride phosphorylase; pH 2.5 acid
phosphatase; periplasmic"
/protein_id="AAC74065.1"
/db_xref="GI:1787215"

```

/translation="KAILIPPLSLILPVTQSAFAQSEPEIKLESVIVYSHGVAP
 TKATQMODVTPDAMPVTPWKIGMLTPPGGELIAYLGHYQORQLVADGLLAKKQDPS
 GOVAIIADVDERTKTEAFAGALPCDAIVHTQADYSSPDLFNPPLTKTGCCQNDNA
 NVTDAIISRAGSIAIDFTGHTQAFRELELVNFPQSNICLKREKODESSCLTOLAPS
 ELKYSADNVSLTGAASLASMLTEIFLQAOGMPEMGCRITDSHOMVTLISLHNAOF
 YLQRTPEVARSRAIPLDLIKTALTATPPPOKATGVTLPTSVLFTAGHDNVLNAG
 ALELWMLPGQPDNTPGCGELVEFERMRSLDSQNIQVSLVFOQLQOMKDKPTSLNT
 PPGEVKTLPGCEENMOGMSLAFQIVNPARIPACSL"
 1153. .1189
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 repeat_region
 1491.1524
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 gene
 complement(1536. .3716)
 /gene="yccc"
 /note="b0981"
 complement(1536. .3716)
 CDS
 complement(1536. .3716)
 /gene="yccc"
 /function="orf; unknown"
 /note="f726; 100 pct identical to fragment yccc_ECOLI
 SM:P38134 but has 627 additional N-terminal aa and 2 less
 C-terminal residues"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical protein"
 /protein_id="AAC74066.1"
 /db_xref="GI:187216"
 /translation="MTKNMTPGSGTOENETDLRLVGLMDHRRKFIISVATFLI
 AVASLSTPIYQADTIYQVQKQKQNALISGSDMIPNPSPEAPETIOLQSRMLLK
 TIAELNDIYQKFTPIVGRMARITKREPELILSMHILPOLNGODQDLTVLIGEN
 GHYTLGESEFTVNGVQGLRLEKQVALITADIKAPGTFVLSQRTLEALNAALDEF
 TVSERKESGMLELMTGDDPOLITRIINLSANNYLNIAQAOSSDSEFTQROL
 PEVSELDQAEKLVNQRQSDVDLNEAKAVLEQIVNVDQLNELFFERELEISOLY
 KKHDTYRALLERKOTLEOERELRNKRVASMSPTQOEVLRLSRDVGAGVYARMLR
 QOELISKSSAIGNRIIDPAVTOPOPKKALNVVIGLILGLFISGAGVYARMLR
 RGVAPRQLEHGISVYATIPMSSEWIDKTRIRKKNLSNOCRRHTKINPIAVNPDA
 DSAVAVALKRSLHFAMETENYILMTGTAPDSGKTFVSTLAAVAQSQKVLFI
 DADLRGYSNHLFTVSNHGLSEYLAGDELKRYIQHKGKGFDTITGQVPPQSEI
 LMRDMLQELMANDHYLDIVDPPLMAVSPAAYGVRSGVSTLLVAFGLTAKREVS
 LSMORLEQAGVNIKGAILINGYIKRASTAVSYGNYYSYSSEKE"
 complement(3733. .3764)
 /note="factor Sigma70; predicted +1 start at 1043443"
 promoter
 complement(3736. .4194)
 /gene="yccc"
 /note="b0982"
 complement(3736. .4194)
 CDS
 complement(3736. .4194)
 /gene="yccc"
 /note="b0982"
 complement(3736. .4194)
 /function="putative enzyme; Not classified"
 /note="f152; This 152 aa ORF is 33 pct identical (13 gaps)
 to 139 residues of an approx. 160 aa protein PPA1_SCHPO
 SM: P41893"
 /codon_start=1
 /transl_table=11
 /product="putative phosphatase"
 /protein_id="AAC74067.1"
 /db_xref="GI:187217"
 /translation="MDSNAQLKFNLSILVCTGNICRSPIGERLLRKRLPGVKYSAG
 VHGVKHPADPADVADYANHGVSLEGHAGARKLTAMARNYDILILMESEHILQVYATA
 PEVKGKTMFLFGOMLEQKEIPDPYRKSQDAFHHVYGMLEKRADEMAKKRLSR"
 complement(3793. .3820)
 promoter
 complement(3793. .3820)
 /gene="yccc"
 /note="factor Sigma70; predicted +1 start at 1043503"
 promoter
 complement(3818. .3844)
 /gene="yccc"
 /note="factor Sigma70; predicted +1 start at 1043528"
 promoter
 complement(3865. .3892)
 /gene="yccc"
 /note="factor Sigma70; predicted +1 start at 1043575"
 gene
 complement(4170. .5309)
 /note="yccc2"
 /note="b0983"
 complement(4170. .5309)
 CDS
 complement(4170. .5309)

/gene="yccc2"
 /function="phenotype; Not classified"
 /note="f379; This 379 aa ORF is 29 pct identical (33 gaps)
 to 364 residues of an approx. 400 aa protein BEXD_HAEIN
 SM: P22236"
 /codon_start=1
 /transl_table=11
 /product="putative function in exopolysaccharide
 production"
 /protein_id="AAC74068.1"
 /db_xref="GI:187218"
 /translation="MKKNIKFSVLTAVLSTACTVPGQNLSTSKNDVTELPNOY
 DLDKWNITPVTPGLIDQIRAPRIMSQAPELEQDIANTERIGIDVLMVWHPPE
 LTPAGQYRSASADTGNWNVADAIFFPYIGRLKGLTQVRNBITAALDSVSPQ
 VDVSAAFRSOKAYVTGVEYSKGOQPIITPLTINDAIIAAGLTADADMRNVLTQ
 GVKTVNLYALMQRGDLRQNLKHPDGLFIRNDLKFVWGVGKOSTLKMDSGM
 TLAEALGNAGENQDQADATGTFVIRATQONKGTANTYOLNARDASAMILGTEFOL
 EPIDIVYVTPARLARNRYSLIVPTISGVHDLTETSRNIQTWPN"
 complement(5355. .7451)
 gene
 complement(5355. .7451)
 /gene="ymca"
 /note="b0984"
 complement(5355. .7451)
 /gene="ymca"
 /note="b0984"
 complement(5355. .7451)
 CDS
 complement(5355. .7451)
 /gene="ymca"
 /note="b0984"
 complement(5355. .7451)
 /function="orf; unknown"
 /note="f698; This 698 aa ORF is 64 pct identical (7 gaps)
 to 694 residues of an approx. 704 aa protein yJRH_ECOLI
 SM: P32689"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical protein"
 /protein_id="AAC74069.1"
 /db_xref="GI:187219"
 /translation="MKKSYLRSCTAIVSSACHAEVLTYPDPILGSSQSDFGGTGLQ
 MNARLAEQESVNV RNDQYRFSTVALEPMLGRTYRDVTRKYSQWEDSGD
 OSYKRSPEFKRLMEGVKWLQVAFGRKIDGTFGLDEYTVAAKQAGPFTLQMA
 KQYANAGNITNTPCRVSDKTYCHRAHSDAGDISDIFRGAASIFGLEYOTPNPPL

Query Match 97.3%; Score 1287.8; DB 1; Length 10784;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 1 ATGAAAGCGATCTTAATCCATTTTATCTCTTGATTCGCTTAACCCCGGAATCTGCA 60
 123 ATGAAAGCGATCTTAATCCATTTTATCTCTTGATTCGCTTAACCCCGGAATCTGCA 182
 61 TTCGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTCGCTCAGTGCATGCT 120
 183 TTCGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGATTCGCTCAGTGCATGCT 242
 121 GTGCGTCTCAACCAAGGCCACGACATGATGACAGATGTCAACCCAGCATGGCCA 180
 243 GTGCGTCTCAACCAAGGCCACGACATGATGACAGATGTCAACCCAGCATGGCCA 302
 181 ACCGTGGCGGTAAACTGGGCTTGCGTGAACACGCGGCGGAGCTATATGCGCTATCTC 240
 303 ACCGTGGCGGTAAACTGGGCTTGCGTGAACACGCGGCGGAGCTATATGCGCTATCTC 362
 241 GGACATTAACCAAGCGGAGCTGTGAGCGGAGATTGCTGGCAAAAAGGCGTCCCG 300
 363 GGACATTAACCAAGCGGAGCTGTGAGCGGAGATTGCTGGCAAAAAGGCGTCCCG 422
 301 CAGTGTGTCAGGTGCGCATTTATTCGTATGTCGACGACGCGTAAACCCGTAACGCGCA 360
 423 CAGTGTGTCAGGTGCGCATTTATTCGTATGTCGACGACGCGTAAACCCGTAACGCGCA 482
 361 GCGTTCGCGCGGCGGCTGGCACCTGACGTGCAATTAACGCTATACCCAGCAGATGCG 420
 483 GCGTTCGCGCGGCGGCTGGCACCTGACGTGCAATTAACGCTATACCCAGCAGATGCG 542
 421 TCACATCCGATCCGCTTATTAATCTCTTAACCAAGTGGCGTGTGCCAATGATGAACGCG 480
 543 TCACATCCGATCCGCTTATTAATCTCTTAACCAAGTGGCGTGTGCCAATGATGAACGCG 602

Oy		481	AACGAGATCGACGGCATCTCAGACAGGCGACGAGGTCATAATTGGTACTTTACCGGGCAT	540
Db		603	AACGTGACTGCACCGGATCTCTACGACAGGCGACGAGGGTCAATTTGCTGACTTTACCGGGCAT	662
Oy		541	CGGCCAAACGGCGGTTTCGGGAACCTGGAACGGGGTGCTTAATTTCCGCATCAAACTTGTGC	600
Db		663	CGGCCAAACGGCGGTTTCGGGAACCTGGAACGGGGTGCTTAATTTCCGCATCAAACTTGTGC	722
Oy		601	CTTAAAGGTGAGAAACAGGACGAAAAGCTGTTCAATTAAACGACGGCATTTACCATCGGAATC	660
Db		723	CTTAAAGGTGAGAAACAGGACGAAAAGCTGTTCAATTAAACGACGGCATTTACCATCGGAATC	782
Oy		661	AAGGTGAGCGCGGACAATGTCCTCAATTAACCGGTCGGGTAAGCCCGCATCAATGCTGACG	720
Db		783	AAGGTGAGCGCGGACAATGTCCTCAATTAACCGGTCGGGTAAGCCCGCATCAATGCTGACG	842
Oy		721	GAGATATTTCTCTCTGCAACAGCACAGGGAATGCCGAGCCGGGGTGGGGAAGATCAACC	780
Db		843	GAGATATTTCTCTCTGCAACAGCACAGGGAATGCCGAGCCGGGGTGGGGAAGATCAACC	902
Oy		781	GATTTCACACCACTGTCGGAACACCTTCTGTAAGTTTGCATTAACGGCAATTTTATTCTTACAA	840
Db		903	GATTTCACACCACTGTCGGAACACCTTCTGTAAGTTTGCATTAACGGCAATTTTATTCTTACAA	962
Oy		841	CGCACGCCAGAGGTTGGCCCGCAGCCGCCACCCCCTAATGATGATTTGATCATGACGACGC	900
Db		963	CGCACGCCAGAGGTTGGCCCGCAGCCGCCACCCCCTAATGATGATTTGATCATGACGACGC	1022
Oy		901	TTCAGCGCCCCATCCACCGCAAAAAACAGGCGGTATGGTGTGACATTAACCCACTTCAGTACTG	960
Db		1023	TTCAGCGCCCCATCCACCGCAAAAAACAGGCGGTATGGTGTGACATTAACCCACTTCAGTACTG	1082
Oy		961	TTTTATTCGCGGACACGATACTTAATCTGGCAAATCTCGGCGGCCGACATGAGACTCAACTGG	1020
Db		1083	TTTTATTCGCGGACACGATACTTAATCTGGCAAATCTCGGCGGCCGACATGAGACTCAACTGG	1142
Oy		1021	ACGCGTCCCGGTCACGCGGATPAACACGCGCCGACGCTGTGAATCGGTGTTGAACGCGTGG	1080
Db		1143	ACGCGTCCCGGTCACGCGGATPAACACGCGCCGACGCTGTGAATCGGTGTTGAACGCGTGG	1202
Oy		1081	CGTGCGGTACGCGATPAACAGCCAGTTCGATTCAGTTCGCTGCTTCACACACTTTCACAG	1140
Db		1203	CGTGCGGTACGCGATPAACAGCCAGTTCGATTCAGTTCGCTGCTTCACACACTTTCACAG	1262
Oy		1141	CAGATGCGTGTAAACCGCCGCTGTCATTAATACGCCGCCGAGAGAGTTAAACTGACC	1200
Db		1263	CAGATGCGTGTAAACCGCCGCTGTCATTAATACGCCGCCGAGAGAGTTAAACTGACC	1322
Oy		1201	CTGGCAGGATGTGAAAGCGGAAAGCGCAGGCGCANCTGCTTGTTGGCAGAGTTTAAACGAA	1260
Db		1323	CTGGCAGGATGTGAAAGCGGAAAGCGCAGGCGCANCTGCTTGTTGGCAGAGTTTAAACGAA	1382
Oy		1261	ATCGTGATGTAAGCACGCGATACCGGCGTGCAGTTTGGAGAT	1300
Db		1383	ATCGTGATGTAAGCACGCGATACCGGCGTGCAGTTTGGAT	1422
RESULT 10				
LOCUS	D90735	18882 bp	DNA	linear
DEFINITION	Escherichia coli genomic DNA. (22.3 - 22.7 min).			BCT 07-FEB-1999
VERSION	D90735.1 GI:4062539			
KEYWORDS	Complete and shotgun sequencing; hyaB; hyaC; hyaD; hyaE; hyaF; appC; cysK; cbdA; appB; cyxB; cbdB; appA; amsI; insA; insB; cpsF; cpsB; sfa; yccL; ycmM.			
SOURCE ORGANISM	Escherichia coli(strain:K12) DNA, clone:kohara clone #225. Escherichia coli Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (sites)			
AUTHORS	Oshima,T., Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Ikemoto,K., Inada,T., Itoh,T., Kajihara,M., Kanai,K., Kashimoto,K.,			

```

TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
  Kimura,S., Kitagawa,M., Makino,K., Masuda,S., Miki,T.,
  Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H.,
  Nishio,Y., Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K.,
  Wada,C., Yamamoto,Y., Yano,M. and Horiiuchi,T.
  A 718-kb DNA sequence of the Escherichia coli K-12 genome
  corresponding to the 12.7-28.0 min region on the linkage map
  DNA Res. 3 (3), 137-155 (1996)
  97061202
2 (sites)
  Alba,H., Babo,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T.,
  Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
  Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,
  Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H.,
  Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T.,
  Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
  Yamamoto,Y. and Yano,M.
  The systematic sequencing of the Escherichia coli genome in Japan
  Unpublished
  3 (bases 1 to 18882)
  Mori,H.
  Direct Submission
  Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science
  and Technology, Res. & Edu. Center for Genetic Info.; 8916-5
  Takayama, Ikoma, Nara 630-01, Japan
  (E-mail:hmori@gtc.aist-nara.ac.jp; Tel:81-7437-2-5660,
  Fax:81-7437-2-5669)
  On Dec 24, 1998 this sequence version replaced gi:1651473.
  Collaboration Information:
  Project:
  The Japan E.coli genome DNA sequencing project
  Group:
  The Japan E.coli genome DNA sequencing group
  Members: (1995.4 - 1996.3)
  Alba,H., Bada,T., Fujita,K., Hayashi,K., Honjo,A.,
  Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S.,
  Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
  Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
  Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K.,
  Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,N.,
  Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
  Yamamoto,Y. and Yano,M.
  Headed by:
  Name: Takashi Horiiuchi
  Address: National Institute of Basic Biology, Okazaki, 444, Japan
  E-mail: kishio@nibb.ac.jp
  Information operator:
  Name: Hirotsada Mori
  Address: NARA Institute of Science and Technology,
  Ikoma, 630-01, Japan
  E-mail: hmori@gtc.aist-nara.ac.jp
  URL:
  The Japan E. coli genome database
  http://sw3.aist-nara.ac.jp.
  Location/Qualifiers
    1..18882
      /organism="Escherichia coli"
      /strain="K12"
      /db_xref="taxon:562"
      /map="22.3-22.7 min"
      /clone="Kohara clone #225"
      /note="Nucleotide position 1034630-1053511 from the
      initiation site of TNA (0 min.).-This clone is from
      Kohara lambda miniset library."
      1..966
        /gene="hyab"
        <1..966
          /gene="hyab"
          /note="ORF_ID:0224#6
          similar to SwissProt Accession Number P19927"
          /codon_start=1
          /transl_table=11
          /product="Hydrogenase-1 large chain (EC 1.18.99.1) (nife
          hydrogenase) (membrane-bound hydrogenase 1 large subunit)

```

```

(hyd1)."
/protein_id="BAA35738.1"
/db_xref="GI:4062540"
/translation="PDALAIGOFNKPMSIEIGTGLSDKCVLSGAPDIAANDGEKSL
MPGAVINGDNNVLPDVLVDPOVDFEYVADHAMVTRPNDQYGRHPFDITDWTIPGD
VKGSTNTIOQLEDERYSWIKAPRWGRNAMEGLPRLIAYHKGDATVSEKWDMS
ALNLPISQISTQILICRAHEAOAAACKLOVFFDKLMTNLSKNGMLATASEKMPAT
WPTFCRGVGTETAPRGALGHMAAIKDGKIDLYOCVPTTWNASPRDPKGOIGAYEAL
MNTKAIEPEOLFELRLTSHSFDPLCACTHYLGDDGSELISQVR"
885. 1692
/gene="hyac"
985. 1692
/gene="hyac"
/note="ORF_ID:0225#1
similar to PIR Accession Number JY0074"
/codon_start=1
/transl_table=11
/product="hyac protein"
/protein_id="BAA35739.1"
/db_xref="GI:1651475"
/translation="MOOKSDNVSHYVEAPVRIWMLTVLCMAVLMTGYEIGKPLP
SVSEATLYFYNGYRIHLHFSAWFTVVLMLRIYAVGNRSRELFIVPWKRSW
OGVWEIRWYFLFKRPSADIGHNPIAOAMPGYLMAVGMITIGSEFALYSOYATF
APRRYVEFEYFGKNSMDIHSWRLGKMLIGAFYIGHYVALREDIMSDTVISTMV
NGIRSHKFKGISNERS"
1689. 2276
/gene="hyad"
1689. 2276
/gene="hyad"
/note="ORF_ID:0225#2
similar to SwissProt Accession Number P19930"
/codon_start=1
/transl_table=11
/product="Hydrogenase-1 operon protein hyad."
/protein_id="BAA35740.1"
/db_xref="GI:1651476"
/translation="MSBORVVMGLGNLMLADDEGCVAVERLIAHYHMPVEYELVDG
GTQGLNLGYVESASHLIDAIDGLDEGTLRTYAGERIPAYISAKKMSLQNSFE
VLALDINGHLPALHAIVLQIPAMLDGEGSLSELAREQLPAAEALAOALAMGIVP
QPNESRCINDCSLMENYEGVRLRQYRMTOEQG"
2273. 2671
/gene="hyae"
2273. 2671
/gene="hyae"
/note="ORF_ID:0225#3
similar to SwissProt Accession Number P19931"
/codon_start=1
/transl_table=11
/product="Hydrogenase-1 operon protein hyae."
/protein_id="BAA35741.1"
/db_xref="GI:1651477"
/translation="MSNDTPDALMQRLMANGWTPVSESRLLDMITQAPDGVLLSSD
PKRTPVEDNPNVIGELRPPDYTMQVAIDLBOSEALIGDRFGVFRPAPATLVFTGN
YRGVLNGIHMAELINLMRGLVEPQOEKAS"
2668. 3525
/gene="hyar"
2668. 3525
/gene="hyar"
/note="ORF_ID:0225#4
similar to SwissProt Accession Number P19932"
/codon_start=1
/transl_table=11
/product="Hydrogenase-1 operon protein hyar."
/protein_id="BAA35742.1"
/db_xref="GI:1651478"
/translation="MSETFHLLRGVTPQNDSPSMPLPTICOVNDEPSMAALEQCA
HSPOVALINELQOLSERQPLGSEVLAVDLINLADNRHITNLLGGEYSVRIQQA
DSESEIOEALFCGLMVRRRRGEKLEDKLEAGCAPALAOAONLLPTDSLLPP
IDGLMNGLPILAHELLAHVRNDAQPHSINKLQPLISEADRLFLSLCGPNIQITIG
YGESYINATGRLHWHLRCDTLKGPLLESEYICPIPEVLAAPEDLVDSQRISEVC
OMLAAPY"
3659. 5203
/gene="appc"
3659. 5203
/gene="appc"

```

```

CDS
/note="cbda; cyxa"
3659. 5203
/gene="appc"
/note="ORF_ID:0225#5
similar to SwissProt Accession Number P26459"
/codon_start=1
/transl_table=11
/product="Cytochrome bd-11 oxidase subunit 1 (EC
1.10.3.-)."
/protein_id="BAA35743.1"
/db_xref="GI:1651479"
/translation="MWDVIDSRMQFALTYHLEFVPLTGLILFLAIMETIYVYG
RTIYDMTRFMWKLFEINFDLGVATGLTMEFQFNSFYSNVDIGAPLAMEALM
AFLESTFVGLFEFGMORLNKQHLVLMVAFGNSLSALMLNANGMWOYPTGAHFD
IDTLREMTSESELVFNPSQYKFEHVTAGVGAEMTAMISAMYLLRGERNVALL
SDAIGSVFTGLAIIGTLOGDSSEAYEVNOVVKLAAMEGQTEPAPFHVWAPF
OOAERNAPAKRIPIALIGTLATSLDKPVGKNTMAETVPRLOGRMAMLMQETISOG
NREPHVIOAFRLGEDDLGYMLISRYAPDMNVYTAQYQAAARCAIPOVAPYNSFRI
MVGCSLLVLMVLAIVOTLRKIDQHRMVLKMLMSLPLPAIATEACMPFTEGRQF
MAIDILPTYSASHSLTTGOLAFSLIMIVGLYTFLLIAYEYLMOKYARLGPAMQSEQ
PTQOOG"
5215. 6351
/gene="appb"
/note="cbdb; cyxb"
5215. 6351
/gene="appb"
/note="ORF_ID:0225#6
similar to SwissProt Accession Number P26458"
/codon_start=1
/transl_table=11
/product="Cytochrome bd-11 oxidase subunit 11 (EC
1.10.3.-)."
/protein_id="BAA35744.1"
/db_xref="GI:1651480"
/translation="MEDYETLRFITWMLIGVILVYFMISDGPDMIGCLLPVARNDD
ERRIYNSVGAHMEGNQWMLIAGALTAANPRVYAAAFSGFYAMITVLSIFRPL
AFDYRKIADARWRKMDAGLVISGVPPVFGINFGNLLGVPAPFQRLREYLLS
FMQLLTPPEPLCGLLSLGMNVLQGGVWLQLTQVGIHLRSQLATRAALLWLCFLA
GYMLWVDIGDVFVLQADNGSPNPLMKLVAVLPGAMNNFVSEPVLMTFPLGFECPPL
LTVMAIYGRGRMGVFLMASLMQFVITFAGITLFFVMPSSVPSISLTLMDSSTSQL
Query Match 97.3%; Score 1287.8; DB 1; Length 18882;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
1 ATGAAGCGATCTATATCCATTTTATCTCTTGATTCCTTAACCCGCAATCTGCA 60
6536 ATGAAGCGATCTTATATCCATTTTATCTCTTGATTCCTTAACCCGCAATCTGCA 6595
61 TTGCTCAGAGTAGACCCGAGCTGAAGCTGGAAGTGTGGTATTCAGTCGTCATGCT 120
6596 TTGCTCAGAGTAGACCCGAGCTGAAGCTGGAAGTGTGGTATTCAGTCGTCATGCT 6655
121 GTGCGTGTCCAAACCAAGCCGACGACGATGAGAGTGTGACCCGACGACGATGGGCA 180
6656 GTGCGTGTCCAAACCAAGCCGACGACGATGAGAGTGTGACCCGACGACGATGGGCA 6715
181 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGNGSTGTGACTATATGCTATCTC 240
6716 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGNGSTGTGACTATATGCTATCTC 6775
241 GGACATTACCAAGCCGACGCTGTGATGCCGACGATTCGTGGCAAAAAGGCTGCCCG 300
6776 GGACATTACCAAGCCGACGCTGTGATGCCGACGATTCGTGGCAAAAAGGCTGCCCG 6835
301 CAGTGTGTCAGTGTGCGCATTTATCTCATATGCGACGAGCGGATACCCGTAACAGGCAA 360
6836 CAGTGTGTCAGTGTGCGCATTTATCTCATATGCGACGAGCGGATACCCGTAACAGGCAA 6895
361 GCCTTCGCGCGCGGCTGGACACTGATGCAATAACCGTACATACCCAGGACAGATACG 420
6896 GCCTTCGCGCGCGGCTGGACACTGATGCAATAACCGTACATACCCAGGACAGATACG 6955

```

```

QY 421 TCAGTCCGACCGTATTATTCCTCTTAAACATGCGCTTTCCTCAACTGATACGCG 480
D 6956 TCAGTCCGACCGTATTATTCCTCTTAAACATGCGCTTTCCTCAACTGATACGCG 7015
QY 481 AACGTACTGACGGATCTCAGACGAGGAGGTCATTTGCTGACTTTACGGGCAT 540
D 7016 AACGTACTGACGGATCTCAGACGAGGAGGTCATTTGCTGACTTTACGGGCAT 7075
QY 541 CGGCAACGCGGTTTCGCAACTGGAACGGGTGCTTATTTTCGCAATCAACTTGGC 600
D 7076 CGGCAACGCGGTTTCGCAACTGGAACGGGTGCTTATTTTCGCAATCAACTTGGC 7135
QY 601 CTTAAACGTAGAAACAGACGAAAGCTGTTTCAATTAACGACGATTAACATCGGAATC 660
D 7136 CTTAAACGTAGAAACAGACGAAAGCTGTTTCAATTAACGACGATTAACATCGGAATC 7195
QY 661 AAGGTAGCGCCGCAATGTCCTATTAAACGGGTGCTTATTTTCGCAATCAACTTGGC 720
D 7196 AAGGTAGCGCCGCAATGTCCTATTAAACGGGTGCTTATTTTCGCAATCAACTTGGC 7255
QY 721 GAGATATTTCTCCGCAACAGACGAAAGCTGGAACGGGTGCTTATTTTCGCAATCAACTTGGC 780
D 7256 GAGATATTTCTCCGCAACAGACGAAAGCTGGAACGGGTGCTTATTTTCGCAATCAACTTGGC 7315
QY 781 GATTCACACAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
D 7316 GATTCACACAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 7375
QY 841 CGCACGCGAAGGTTGCCCGACGCGGCAACCCCGTTATTTGATTTGATGATGCGACG 900
D 7376 CGCACGCGAAGGTTGCCCGACGCGGCAACCCCGTTATTTGATTTGATGATGCGACG 7435
QY 901 TTGACGCGCCATCCACGCGAAGGCGGATGCTGATGATTAACCATCTTCACTGACTG 960
D 7436 TTGACGCGCCATCCACGCGAAGGCGGATGCTGATGATTAACCATCTTCACTGACTG 7495
QY 961 TTTATTCGCGACACGATTAATCTGGAATATCTCGCGGCGGCTGAGTGAAGTGAAGTGA 1020
D 7496 TTTATTCGCGACACGATTAATCTGGAATATCTCGCGGCGGCTGAGTGAAGTGAAGTGA 7555
QY 1021 ACGCTTCCGCGTACGCGGATTAACAGCCCGGCGGATGCTGATGATTAACCATCTTCACTGACTG 1080
D 7556 ACGCTTCCGCGTACGCGGATTAACAGCCCGGCGGATGCTGATGATTAACCATCTTCACTGACTG 7615
QY 1081 CGTGGGCTAAGCGATTAACAGCGATTAACAGCGGCTGCTGCTTCCACACTTTCAG 1140
D 7616 CGTGGGCTAAGCGATTAACAGCGATTAACAGCGGCTGCTGCTTCCACACTTTCAG 7675
QY 1141 CAGATGCGTATTAACAGCGGCTGCTTATTAATACGCGCGCGGAGAGGTAACCTGACC 1200
D 7676 CAGATGCGTATTAACAGCGGCTGCTTATTAATACGCGCGCGGAGAGGTAACCTGACC 7735
QY 1201 CTGCGAGATGTAAGAGCGAAATGCGAGGCGATGCTGCTTTCGAGGATTTTACGAA 1260
D 7736 CTGCGAGATGTAAGAGCGAAATGCGAGGCGATGCTGCTTTCGAGGATTTTACGAA 7795
QY 1261 ATGCTGAATGAAGACCGATTAACGCGGCTGCTGAGTGAAGT 1300
D 7796 ATGCTGAATGAAGACCGATTAACGCGGCTGCTGAGTGAAGT 7835

```

```

RESULT 11
ECOCAPHB 1901 bp DNA linear BCT 04-APR-2002
LOCUS 103371
DEFINITION (appa) gene, complete cds.
ACCESSION L03371.1 GI:145178
VERSION 1
KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
SOURCE Escherichia coli.
ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

```

```

REFERENCE 1 (sites)
AUTHORS Dassa,J., March,C. and Boquet,P.L.
TITLE The complete nucleotide sequence of the Escherichia coli gene appa reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616
PUBMED 2168385
REFERENCE 2 (bases 1 to 1901)
AUTHORS Ostanin,K., Harns,E.H., Stevls,P.E., Kuciel,R., Zhou,M.M. and Van Elteren,R.L.
TITLE Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase
JOURNAL J. Biol. Chem. 267 (32), 22830-22836 (1992)
PUBMED 93054596
FEATURES
source location/Qualifiers
1..1901
/organism="Escherichia coli"
/db_xref="taxon:562"
188..1486
/gene="appa"
188..1486
/gene="appa"
/standard_name="acid phosphatase"
/codon_start=1
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AA00003.1"
/db_xref="GI:145179"
/translation="MKAILIPFLSLIPFPOSAPQSEPELKLESVIVSRNGBRAP
TKATQIMQDVPDAMPMPYKLTGLTRGSLTAYLGHYQORLYANGLLAKKCPROS
GQVALIDVDERTKRGARAGAPDALTIVHPQDSPLPPLKTCGQLDS
NVTALISIRAGSIADEFGHROTAFRELEVLNTPQSLCKRKQDESCSLTGLDS
ELKVASDNVSLTGAIVSLASMLTEIFLQAOAGPEPGMGRITDSHOWNTLISNAQF
YLQRTPEVARSRAPIPLDLKLTALTPHPQKQVGYLTPSVLFIAGHDNLNLG
ALELNTLPQGPDPGGLVFEFRMRSLSPNSOMIOVSIYFQTLQMRDXTPLSLMT
PRGEKTLIACGERBNMGMSLQFQIVNHAIRPACSL"
variation
302..304
/gene="appa"
/citation=[1]
/phenotype="H17N mutant"
BASE COUNT 475 a 498 c 499 g 428 t 1 others
ORIGIN
Query Match 97.2%; Score 1286.2; DB 1; Length 1901;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

361 GCCTTGGCCGCGGGCTGGCACCTGCTGCAATTAACCGATACCATCCAGGACAGATACG 420
 548 GCCTTGGCCGCGGGCTGGCACCTGCTGCAATTAACCGATACCATCCAGGACAGATACG 607
 421 TCCAGTCCCGGATCCCTTATTAATCCCTTAATAAATGGCGTTTGGCAACTGGATTAACGG 480
 608 TCCAGTCCCGGATCCCTTATTAATCCCTTAATAAATGGCGTTTGGCAACTGGATTAACGG 667
 481 AACGTGATGACGCGATCTCAGCAGGAGGAGGATTAATGCTGATTTACCGGGCAT 540
 668 AACGTGATGACGCGATCTCAGCAGGAGGAGGATTAATGCTGATTTACCGGGCAT 727
 541 CGGCAACGCGGTTTCCGGAATGAGACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
 728 CGGCAACGCGGTTTCCGGAATGAGACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 787
 601 CTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATTTACCATCCGGAATC 660
 788 CTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATTTACCATCCGGAATC 847
 661 AAGGTGACGCGGATCTCTCATTAACCGGTGGGTAAAGCTTCGATTCGATTCGATGAG 720
 848 AAGGTGACGCGGATCTCTCATTAACCGGTGGGTAAAGCTTCGATTCGATTCGATGAG 907
 721 GAGATATTTCTCTGCAACAAGACAGGGAATCCGGAACCGGGGTGGGAAAGATCAC 780
 908 GAGATATTTCTCTGCAACAAGACAGGGAATCCGGAACCGGGGTGGGAAAGATCAC 967
 781 GATTCACACAGAGTGAACACCTGCTGATTAACGCGGATTTGATTAATTTTGGTACAA 840
 968 GATTCACACAGAGTGAACACCTGCTGATTAACGCGGATTTGATTAATTTTGGTACAA 1027
 841 CGCAGCCGAGAGTTGCCGAGCGCCGACCCCGATTATGATTTGATCATGAGCAGCG 900
 1028 CGCAGCCGAGAGTTGCCGAGCGCCGACCCCGATTATGATTTGATCATGAGCAGCG 1087
 901 TTGAGCGCCCATCCACCCGCAAAACAGCGGATGATGATGATTAACCATTTAGTACTG 960
 1088 TTGAGCGCCCATCCACCCGCAAAACAGCGGATGATGATGATTAACCATTTAGTACTG 1147
 961 TTTATGCGGAGACAGTACTAATCTGCAATCTCGGCGGAGGAGTGGAGTCAACTGAG 1020
 1148 TTTATGCGGAGACAGTACTAATCTGCAATCTCGGCGGAGGAGTGGAGTCAACTGAG 1207
 1021 ACGCTTCCCGGTACCGGATTAACAGCGCGGAGGATGATGATGATTTGAAAGCTGAG 1080
 1208 ACGCTTCCCGGTACCGGATTAACAGCGCGGAGGATGATGATGATTTGAAAGCTGAG 1267
 1081 CGTGGCGCTAAGCGATTAACAGCGGATGATGATGATGATGATTTGAAAGCTGAG 1140
 1268 CGTGGCGCTAAGCGATTAACAGCGGATGATGATGATGATGATTTGAAAGCTGAG 1327
 1141 CAGATGCGGTGATTAACAGCGGATGATGATGATGATGATTTGAAAGCTGAG 1200
 1328 CAGATGCGGTGATTAACAGCGGATGATGATGATGATGATTTGAAAGCTGAG 1387
 1201 CTGGCAGGATGAGAGCGAAATGCGCAGGAGGATGATGATGATTTGAAAGCTGAG 1260
 1388 CTGGCAGGATGAGAGCGAAATGCGCAGGAGGATGATGATGATTTGAAAGCTGAG 1447
 1261 ATCGTGAATGAAGCAGCATACCGGCTGACGTTTGACAT 1300
 1448 ATCGTGAATGAAGCAGCATACCGGCTGACGTTTGACAT 1487

SOURCE Escherichia coli.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE 1
 AUTHORS Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.
 TITLE Recombinant Bacterial Phylases and Uses Thereof
 JOURNAL Patent: WO 0190333-A 5 29-NOV-2001;
 DIVERSA CORPORATION (US)
 FEATURES location/Qualifiers
 source 1..1901
 BASE COUNT 475 a 499 c 498 g 428 t 1 others
 ORIGIN
 Query Match 97.2%; Score 1286.2; DB 6; Length 1901;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 1 ATGAAGCGATCTTATATCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
 188 ATGAAGCGATCTTATATCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 247
 61 TTGCTCAGATGACGCGGAGCTGAGCTGGAAGTGTGATTTGATGCTGATGCT 120
 248 TTGCTCAGATGACGCGGAGCTGAGCTGGAAGTGTGATTTGATGCTGATGCT 307
 121 GTGGCTGCTCAACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 308 GTGGCTGCTCAACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 367
 181 ACCTGCGCGGTAAACCTGGTGTGAGTACGACGCGGCGGTGTGAGTACGCTATCTC 240
 368 ACCTGCGCGGTAAACCTGGTGTGAGTACGACGCGGCGGTGTGAGTACGCTATCTC 427
 241 GGACATTAACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 428 GGACATTAACCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 487
 301 CAGTGTGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
 488 CAGTGTGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 547
 361 GCCTTGGCCGCGGGCTGGCACCTGCTGCAATTAACCGATACCATCCAGGACAGATACG 420
 548 GCCTTGGCCGCGGGCTGGCACCTGCTGCAATTAACCGATACCATCCAGGACAGATACG 607
 421 TCCAGTCCCGGATCCCTTATTAATCCCTTAATAAATGGCGTTTGGCAACTGGATTAACGG 480
 608 TCCAGTCCCGGATCCCTTATTAATCCCTTAATAAATGGCGTTTGGCAACTGGATTAACGG 667
 481 AACGTGATGACGCGATCTCAGCAGGAGGAGGATTAATGCTGATTTACCGGGCAT 540
 668 AACGTGATGACGCGATCTCAGCAGGAGGAGGATTAATGCTGATTTACCGGGCAT 727
 541 CGGCAACGCGGTTTCCGGAATGAGACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
 728 CGGCAACGCGGTTTCCGGAATGAGACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 787
 601 CTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATTTACCATCCGGAATC 660
 788 CTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATTTACCATCCGGAATC 847
 661 AAGGTGACGCGGATCTCTCATTAACCGGTGGGTAAAGCTTCGATTCGATTCGATGAG 720
 848 AAGGTGACGCGGATCTCTCATTAACCGGTGGGTAAAGCTTCGATTCGATTCGATGAG 907
 721 GAGATATTTCTCTGCAACAAGACAGGGAATGCGGAGGAGGAGGAGGAGGAGGAGGAG 780
 908 GAGATATTTCTCTGCAACAAGACAGGGAATGCGGAGGAGGAGGAGGAGGAGGAGGAG 967

RESULT 12
 LOCUS AX356570 1901 bp DNA linear PAT 06-FEB-2002
 DEFINITION Sequence 5 from Patent WO0190333.
 ACCESSION AX356570
 VERSION AX356570.1 GI:18621057
 KEYWORDS

OY	781	TAATTCACACCCAGTGGAAACACCTTCTCTAACTTTGATATAACGGCGCAATTTTATTCCTCAAA	840
Db	968	GAATTCACACCAAGTGGAAACACCTTCTCTAACTTTGATATAACGGCGCAATTTTATTCCTCAAA	1027
OY	841	CGCACGGCCAGAGGTGGCCCGCAGCGCGCCACCCCGCTATTTGGATTGGATCATGTGCAAGCG	900
Db	1028	CGCACGGCCAGAGGTGGCCCGCAGCGCGCCACCCCGCTATTTAGATTGGATCATGAACAGCG	1087
OY	901	TTTGACGGCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTATCCGACCTTCAGTACTG	960
Db	1088	TTTGACGGCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTATCCGACCTTCAGTACTG	1147
OY	961	TTTATTGCGCGGACACGATACTAAATCTGGCAATCTCGGCGGCGGCACTGAGACTCAACTG	1020
Db	1148	TTTATTGCGCGGACACGATACTAAATCTGGCAATCTCGGCGGCGGCACTGAGACTCAACTG	1207
OY	1021	ACGCTTCCCGGTACACGGGATTAACACGCCCGCCAGGTGGTGAATCTGGTCTTTGAACGCTGG	1080
Db	1208	ACGCTTCCCGGTACACGGGATTAACACGCCCGCCAGGTGGTGAATCTGGTCTTTGAACGCTGG	1267
OY	1081	CGTGGCGCTAAACGATTAACACGACGACATGTGATTCAGGCTTTCGCGGCTCTTCACAGCTTTACAG	1140
Db	1268	CGTGGCGCTAAACGATTAACACGACGACATGTGATTCAGGCTTTCGCGGCTCTTCACAGCTTTACAG	1327
OY	1141	CAGATGCGTGAATAAAAGCCGCTGTCATTAATAACGCCGCCGAGAGAGGTGAACCTGACC	1200
Db	1328	CAGATGCGTGAATAAAAGCCGCTGTCATTAATAACGCCGCCGAGAGAGGTGAACCTGACC	1387
OY	1201	CTGGCAGGATGTGAAGAGACGAAATAGCGCAGGCGCATGTTCGTTGGCAGGTTTTACGGAA	1260
Db	1388	CTGGCAGGATGTGAAGAGACGAAATAGCGCAGGCGCATGTTCGTTGGCAGGTTTTACGGAA	1447
OY	1261	ATCGTGAATGAAGACACGCATACCGGCGGTGCACGTTTGAAGAT	1300
Db	1448	ATCGTGAATGAAGACACGCATACCGGCGGTGCACGTTTGAAGAT	1487

RESULT 13					
AX356574					
LOCUS	AX356574 .	1901 bp	DNA	linear	PAT 06-FEB-2002D
DEFINITION	Sequence 9 from patent WO0190333.				
ACCESSION	AX356574				
VERSION	AX356574.1	GI:18621060			
KEYWORDS					
SOURCE					
ORGANISM	Escherichia coli. Escherichia coli Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
REFERENCE	1 Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and O'Donoghue,E. Recombinant bacterial phycases and uses thereof Patent: WO 0190333-A 9 29-NOV-2001;				
TITLE	JOURNAL				
FEATURES	DIVERSA CORPORATION (us) Location/Qualifiers 1..1901 /organism="Escherichia coli" /db_xref="taxon:562"				
BASE COUNT	473 a	499 c	500 g	428 t	1 others
ORIGIN					
Query Match	97.2%	Score 1286.2;	DB 6;	Length 1901;	
Best Local Similarity	99.4%;	Pred. No. 0;			
Matches 1292:	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;	
OY	1 ATGAAGCGATCTTAATCCATTTTTATCTCTTGATTCGGTTAACCCGCCAAATCTGCA	60			
Db	188 ATGAAGCGACTTAATCCCATTTTATCTCTTGATTCGGTTAACCCGCCAAATCTGCA	247			
OY	61 TTCCTCAGATGAGCCGCAGCTGAGCTGGAAGTGTGGATTTGCATGCTCATGCT	120			
Db	248 TTCCTCAGATGAGCCGCAGCTGAGCTGGAAGTGTGGATTTGCATGCTCATGCT	307			

OY	121	GTGGTGTCCTCAACCAAGGCCAGCACTGATGTGAGATGTGCACCCAGACGATGGCGCA	180
Db	308	GTGGGTGTCTCCAAACCAAGGCCAGCACTGATGTGAGATGTGCACCCAGACGATGGCGCA	367
OY	181	ACCTGGCCGGTAAAACTGGTGTGGCTGACACCCGGNGGTGGTGTGAGCTAATCGCTATCTC	240
Db	368	ACCTGGCCGGTAAAACTGGTGTGGCTGACACCCGGNGGTGGTGTGAGCTAATCGCTATCTC	427
OY	241	GGACATTACCAACGCCACAGCTCTGGTAGCCAGCAGGATGTGCTGGGAAAAAAGGCTGCCCG	300
Db	428	GGACATTATACCAACGCCAGCGCTCTGGTAGCCAGCAGGATGTGCTGGGAAAAAAGGCTGCCCG	487
OY	301	CAGTCTGTGTCAGGTCTCGGATTAATGTCTGATGTCCAGCAGCGTACCCTGTAACAGCGGAA	360
Db	488	CAGTCTGTGTCAGGTCTCGGATTAATGTCTGATGTCCAGCAGCGTACCCTGTAACAGCGGAA	547
OY	361	GCGTTGGCGCGCGGGGGCTGGCAGCGAGCTGTGCATTAACCCGTAACCCGATCCAGCAGATACG	420
Db	548	GCGTTGGCGCGCGGGGGCTGGCAGCGAGCTGTGCATTAACCCGTAACCCGATCCAGCAGATACG	607
OY	421	TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTGTGTGCAACTGGATTAACGCG	480
Db	608	TCCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTGTGTGCAACTGGATTAACGCG	667
OY	481	AACGTGACTGACCGCGATCTCTGACAGAGGCGAGGAGGTCAATTGCTGACTTAACGGGCAT	540
Db	668	AACGTGACTGACCGCGATCTCTGACAGAGGCGAGGAGGTCAATTGCTGACTTAACGGGCAT	727
OY	541	CGGCAAAACGGGGTTTCGGCACTGGAACGGGTGCTTAATTTTCGCAATCAAACTGTGCG	600
Db	728	CGGCAAAACGGGGTTTCGGCACTGGAACGGGTGCTTAATTTTCGCAATCAAACTGTGCG	787
OY	601	CTTAAACGTGAGAAAAACAGAGGAAAGCTGTTCAATTAAACGAGGCAATTACATCGGAATC	660
Db	788	CTTAAACGTGAGAAAAACAGAGGAAAGCTGTTCAATTAAACGAGGCAATTACATCGGAATC	847
OY	661	AAGGTGAGCGCGCGACATATGTCTCATTTAACCGGTGCGGTAAACCTTCGATCAATGCTGACG	720
Db	848	AAGGTGAGCGCGCGACATATGTCTCATTTAACCGGTGCGGTAAACCTTCGATCAATGCTGACG	907
OY	721	GAGATTAATTTCTCGCAACAGCAACAGGAAATGCGGAGCGGGGTGGGGAAGATCAAC	780
Db	908	GAGATTAATTTCTCTGCAACAGCAACAGGAAATGCGGAGCGGGGTGGGGAAGATCAAC	967
OY	781	GATTTCACACACAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA	840
Db	968	GATTTCACACACAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA	1027
OY	841	CGCACGCCAGAGTGGCCCCGCGACGCCGCCACCCGTTATTTGATTTGATCATGCGACGCG	900
Db	1028	CGCACGCCAGAGTGGCCCCGCGACGCCGCCACCCGTTATTTGATTTGATCATGCGACGCG	1087
OY	901	TTTGAGCGCCCATCCACCGCAAAAACAGCGGATATGCTGTGATATACCACTTCAGTACTG	960
Db	1088	TTTGAGCGCCCATCCACCGCAAAAACAGCGGATATGCTGTGATATACCACTTCAGTACTG	1147
OY	961	TTTATTTGCGCGACACGATCTAATCTGGCAATCTCGCGCGCGCACTGAGACTCAACTG	1020
Db	1148	TTTATTTGCGCGACACGATCTAATCTGGCAATCTCGCGCGCGCACTGAGACTCAACTG	1207
OY	1021	ACGCTTCCCGGTACAGCCGGATTAACAGCCGCCAGGTGTGATAGTGTGTTGAACGCTGG	1080
Db	1208	ACGCTTCCCGGTACAGCCGGATTAACAGCCGCCAGGTGTGATAGTGTGTTGAACGCTGG	1267
OY	1081	CGTGGGGCTAAACGATTAACAGCCAGTGGATTAAGTTTGCGTGGCTTCACACTTTPACAG	1140
Db	1268	CGTGGGGCTAAACGATTAACAGCCAGTGGATTAAGTTTGCGTGGCTTCACACTTTPACAG	1327
OY	1141	CAGATGGCGTAAACAGCCGCTGTCTCAATTAATAGCGCGCCCGAGAGAGTAAACTGAC	1200
Db	1328	CAGATGGCGTAAACAGCCGCTGTCTCAATTAATAGCGCGCCCGAGAGAGTAAACTGAC	1387
OY	1201	CTGGCGAGATGTGAAGACGGAATGTCCGAGGGCATGTGTTCGTTGGCAGATTTTACGCCAA	1260

|||||
Db 1388 CCGGAGGATGTGAAGAGCAATGCGGCGCATGTCTGTCGACAGTTTACGCAA 1447
Qy 1261 ATCTGTAATGAAGACGATACCGGCGTGCAGTTTGAGAT 1300
|||||
Db 1448 ATCTGTAATGAAGACGATACCGGCGTGCAGTTTGTAT 1487
RESULT 14
ECOCAPHC 1901 bp DNA linear BCT 04-APR-2002
LOCUS Escherichia coli periplasmic phosphoanhydride phosphohydrolase
DEFINITION (appa) gene, complete cds.
ACCESSION L03372.1 GI:145180
VERSION L03372.1
KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (sites)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
TITLE The complete nucleotide sequence of the Escherichia coli gene appa
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase
J. Bacteriol. 172 (9), 5497-5500 (1990)
JOURNAL 90368616
MEDLINE 2168385
PUBMED
REFERENCE 2 (bases 1 to 1901)
AUTHORS Ostania, K., Harms, E.H., Stevis, P.E., Kuciel, R., Zhou, M.M. and Van
Etten, R.L.
TITLE Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase
J. Biol. Chem. 267 (32), 22830-22836 (1992)
JOURNAL 93054596
MEDLINE 1429631
PUBMED
FEATURES
source
1. Location/Qualifiers
/organism="Escherichia coli"
/db_xref="taxon:562"
188..1486
/gene="appa"
188..1486
/gene="appa"
/standard_name="acid phosphatase"
/citation-[1]
/codon_start=1
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AA00004.1"
/db_xref="GI:145181"
/translation="MKAILIPLSLILPLPOSAPFAQSEPELLESVIVSRHGAAP
TKATOLMDVTPDAPTPVKLGMPLPRGGLIAYIGHYORORLVAADGLAKGPOS
GOVAIIADVDETRKTEGEAFAGLAADCAITVHOADISSPDLFNPILKTGVCOLDA
NVTDAIIRAGSGINDPTGHNPTARELRVINTFOSNICLREKQDESCSLTQALPS
ELKVSADNVSILGAVSLASMLTEIFLLQDQGMPEPGRITDSHOWMTLLSLNAOF
YLELMTPLPGPDNTPGGEVLTFERMRRLSDNSOMIOWSLVYQTLQOMRDKPLSLNT
PPGEVKLTLAGEERNAQMGCSLAFGTQIVNARIPACSL"
299..301
variation
/gene="appa"
/citation-[1]
/phenotype="R20A mutant"
BASE COUNT 474 a 499 c 499 g 428 t 1 others
ORIGIN
Query Match 97.1%; Score 1284.6; DB 1; Length 1901;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTGATTCGTTAAACCGGCAATCGA 60
|||||
Db 188 ATGAAGGAGATCTTAATCCATTTTATCTCTTGATTCGTTAAACCGGCAATCGA 247

Qy 61 TTCGTCAGAGTAGACCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCAAGTCATGCT 120
|||||
Db 248 TTCGTCAGAGTAGACCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCAAGTCATGCT 307
|||||
Qy 121 GTGCGTCTCCCAACCAAGGCCACGCACTGATGACAGATGTACACCCAGACCATGCGCA 180
|||||
Db 308 GTGGCTGCTCCCAACCAAGGCCACGCACTGATGACAGATGTACACCCAGACCATGCGCA 367
|||||
Qy 181 ACCGTGCGGTAAACTGCGGTGTGCTGACACCGCGGCGGTGTGAGCTAATCCCTATCTC 240
|||||
Db 368 ACTGTGCGGTAAACTGCGGTGTGCTGACACCGCGGCGGTGTGAGCTAATCCCTATCTC 427
|||||
Qy 241 GGACATTTACCAACGCCAGCGTGTGTGAGCGGATTTGCGGCAAAAAGGGCTGCGCG 300
|||||
Db 428 GGACATTTACCAACGCCAGCGTGTGTGAGCGGATTTGCGGCAAAAAGGGCTGCGCG 487
|||||
Qy 301 CAGTCTGGTCAAGTGGCGGATTTATGTGATGTGACAGACGATACCCGTAAACAGCGCA 360
|||||
Db 488 CAGTCTGGTCAAGTGGCGGATTTATGTGATGTGACAGACGATACCCGTAAACAGCGCA 547
|||||
Qy 361 GCCTTCGCGCGGCGGCTGCGACCTGACATGCAATTAACCGTATACCGGAGCATAGC 420
|||||
Db 548 GCCTTCGCGCGGCGGCTGCGACCTGACATGCAATTAACCGGAGCATAGC 607
|||||
Qy 421 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGTTTGCACACTGATTAACGCG 480
|||||
Db 608 TCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGGTTTGCACACTGATTAACGCG 667
|||||
Qy 481 AACGTGATGACAGCGCATCTCTGACAGAGGAGGATCAATTTGCTGATTTACCGGCGAT 540
|||||
Db 668 AACGTGATGACAGCGCATCTCTGACAGAGGAGGATCAATTTGCTGATTTACCGGCGAT 727
|||||
Qy 541 CGGCAAAAGCGGCTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGTC 600
|||||
Db 728 CGGCAAAAGCGGCTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGTC 787
|||||
Qy 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGAGCATTTACCATGGAATCT 660
|||||
Db 788 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGAGCATTTACCATGGAATCT 847
|||||
Qy 661 AAGGTGAGCGCGGACAAATGTCATTAACGGTGGGATGAGCTGCAATGCTGAGGAG 720
|||||
Db 848 AAGGTGAGCGCGGACAAATGTCATTAACGGTGGGATGAGCTGCAATGCTGAGGAG 907
|||||
Qy 721 GAGATATTTCTCTCGCAACAGCAAGGAAATCCGCGGAGCGGAGTGGGAAAGATCAAC 780
|||||
Db 908 GAGATATTTCTCTCGCAACAGCAAGGAAATCCGCGGAGCGGAGTGGGAAAGATCAAC 967
|||||
Qy 781 GATTCACACCACTGGAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 840
|||||
Db 968 GATTCACACCACTGGAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 1027
|||||
Qy 841 CGCAGCCGAGAGTGGTCCCGGACGCGGCGCACCCGCTTATGATTTGATCAATGCGAGG 900
|||||
Db 1028 CGCAGCCGAGAGTGGTCCCGGACGCGGCGCACCCGCTTATGATTTGATCAAGAGAGG 1087
|||||
Qy 901 TTGACGCGCCATCCACCGCAAAAGAGCGATGATGTGATCAATTAACCCACTTCACTAGT 960
|||||
Db 1088 TTGACGCGCCATCCACCGCAAAAGAGCGATGATGTGATCAATTAACCCACTTCACTAGT 1147
|||||
Qy 961 TTTATTCGCGGACAGCATTAATCTGCAAAATCTGCGGCGGCACTGGAGCTCAATGCG 1020
|||||
Db 1148 TTTATTCGCGGACAGCATTAATCTGCAAAATCTGCGGCGGCACTGGAGCTCAATGCG 1207
|||||
Qy 1021 ACGCTTCCCGGTACACCGGATTAACACCGCGGAGTGGATGATGCTGTTGAAGCGTGG 1080
|||||
Db 1208 ACGCTTCCCGGTACACCGGATTAACACCGCGGAGTGGATGATGCTGTTGAAGCGTGG 1267
|||||
Qy 1081 CGTGGCTAAGGAGTAACAGCAGTGTGATTCAGTTTGTGCTGCTTCCAGACTTTAAG 1140
|||||
Db 1268 CGTGGCTAAGGAGTAACAGCAGTGTGATTCAGTTTGTGCTGCTTCCAGACTTTAAG 1327

Oy 1141 CAGATGGCTGATAAAGCCGCTGTCAATTAATACGGCCCGGAGAGGTAAATGACC 1200
 |||||||
 Db 1328 CAGATGGCTGATAAAGCCGCTGTCAATTAATACGGCCCGGAGAGGTAAATGACC 1387
 |||||||
 Oy 1201 CTGGCAGAGATGTAAGAGCCGAATATGACGAGGATGTCTGTTGGCAGGTTTACCAA 1260
 |||||||
 Db 1388 CTGGCAGAGATGTAAGAGCCGAATATGACGAGGATGTCTGTTGGCAGGTTTACCAA 1447
 |||||||
 Oy 1261 ATCGTGAATGACGACCATACCGCGCTGCAGTTTGAGAT 1300
 |||||||
 Db 1448 ATCGTGAATGACGACCATACCGCGCTGCAGTTTGAT 1487
 |||||||

RESULT 15
 ECOCAPHD 1901 bp DNA linear BCT 04-Apr-2002
 LOCUS
 DEFINITION Escherichia coli periplasmic phosphoanhydride phosphohydrolase
 (appa) gene, complete cds.
 ACCESSION L03373.1 GI:145182
 VERSION L03373.1
 KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE 1 (sites)
 AUTHORS Dassa, J., Marck, C. and Boquet, P. L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appa
 reveals a significant homology between pH 2.5 acid phosphatase and
 glucose-1-phosphatase
 J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616
 PUBMED 2168385
 2 (bases 1 to 1901)
 Oostan, R., Harms, E. H., Stevis, P. E., Kuciel, R., Zhou, M. M. and Van
 Etten, R. L.
 TITLE Overexpression, site-directed mutagenesis, and mechanism of
 Escherichia coli acid phosphatase
 J. Biol. Chem. 267 (32), 22830-22836 (1992)
 MEDLINE 93054596
 PUBMED 1429631

FEATURES
 source location/Qualifiers
 1..1901
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 188..1486
 /gene="appa"
 /gene="appa"
 /standard_name="acid phosphatase"
 /citation=[1]
 /codon_start=1
 /transl_table=1
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AA00005.1"
 /db_xref="GI:145183"
 /translation="MKAILIPELSLIPITPQSAFQSEPEIKIESVIVSRHGVRA
 TKATIQMDTPDPAHPTWPKIKGLTPRGELIAYLGHQRLVAGLAKKCPDS
 GOVAIADVBERKTEAFAGLAPCAIVTHQADTSFDFNPLKGTGVCGLDNA
 NVTDAISRAAGSIADTGHQRTAFRELERYLNPOSNLCKREKODESCSLNALPS
 ELKVSADNVSLTGAVSLASMLTEIFLQQAQGMPEGMGRITDSHONNTLLSLNAQF
 YLQRTPEVARSRAITPLDLIKTALTPHPQKQAYGVTLPTSVFLIGHDNLNLNG
 ALKNTLTPGQDPNTPPGELVPERMRRLSDNSQWIOVSLVFOILQMRDKTPISLMT
 PRGVKTLTLAGCERNAGKMSLAGFTQIVNARIIPACSL"
 299..301
 /gene="appa"
 /citation=[1]
 /phenotype="R16a mutant"
 BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN

Query Match 97.1%; Score 1284.6; DB 1; Length 1901;
 Best Local Similarity 99.3%; Pred. No. 0;

Matches 1291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Oy 1 ATGAAGGATCTTAATCCATTTTATCTTGTATCTCCGTTAACCCGCAATCTGCA 60
 |||||||
 Db 188 ATGAAGGATCTTAATCCATTTTATCTTGTATCTCCGTTAACCCGCAATCTGCA 247
 |||||||
 Oy 61 TTGCGTAGAGTGGCCGGAGCTAGAGTGAAGTGGTGGATGTCAGTCGTCATAGT 120
 |||||||
 Db 248 TTGCGTAGAGTGGCCGGAGCTAGAGTGAAGTGGTGGATGTCAGTCGTCATAGT 307
 |||||||
 Oy 121 GTGCGTGTCCCAACCAAGGCGCACCACTGATGAGATGTCACCCAGAGCATGGCCA 180
 |||||||
 Db 308 GTGCGTGTCCCAACCAAGGCGCACCACTGATGAGATGTCACCCAGAGCATGGCCA 367
 |||||||
 Oy 181 ACTTGGCCGTTAAACTGAGTGGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 240
 |||||||
 Db 368 ACTTGGCCGTTAAACTGAGTGGCTGACACCGCGNGGTGTGAGCTAATCGCTATCTC 427
 |||||||
 Oy 241 GGACATTAACCAAGCGCGTGTGTAAGCCGAGGATTTGGCGGAAAGGCGTGGCCG 300
 |||||||
 Db 428 GGACATTAACCAAGCGCGTGTGTAAGCCGAGGATTTGGCGGAAAGGCGTGGCCG 487
 |||||||
 Oy 301 CAGTCTGTGATGATGCGGATTAATGCTGATGTCGACGAGCGTACCCGTTAAACAGCGGAA 360
 |||||||
 Db 488 CAGTCTGTGATGATGCGGATTAATGCTGATGTCGACGAGCGTACCCGTTAAACAGCGGAA 547
 |||||||
 Oy 361 GCGTTGCGCGCGCGGCTGGACCTGATGTCGATTAACCGTACATCCAGGCGATACG 420
 |||||||
 Db 548 GCGTTGCGCGCGCGGCTGGACCTGATGTCGATTAACCGTACATCCAGGCGATACG 607
 |||||||
 Oy 421 TCAGTCCGCGATTCGTTATTAATCTTAATAAAGCGGTTGCCAATCGATTAACGCG 480
 |||||||
 Db 608 TCAGTCCGCGATTCGTTATTAATCTTAATAAAGCGGTTGCCAATCGATTAACGCG 667
 |||||||
 Oy 481 AACGTGATGACGCGATCTTCAGCAGGCGCAGAGGCTCAATGCTGACTTAAACCGGCAAT 540
 |||||||
 Db 668 AACGTGATGACGCGATCTTCAGCAGGCGCAGAGGCTCAATGCTGACTTAAACCGGCAAT 727
 |||||||
 Oy 541 CCGCAACGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGCG 600
 |||||||
 Db 728 CCGCAACGCGGTTTCCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGCG 787
 |||||||
 Oy 601 CTTAAAGCTGGAACAGGAGCAAGAGCGTTCATTAACGCGGATTAACATTCATTCGGAATC 660
 |||||||
 Db 788 CTTAAAGCTGGAACAGGAGCAAGAGCGTTCATTAACGCGGATTAACATTCATTCGGAATC 847
 |||||||
 Oy 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 720
 |||||||
 Db 848 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 907
 |||||||
 Oy 721 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAACCGGGGTGGGGAAGATCAC 780
 |||||||
 Db 908 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAACCGGGGTGGGGAAGATCAC 967
 |||||||
 Oy 781 GATTACACACAGTGAACACCTTCTAAGTTTGATTAACGGCAATTTTATTTCTCTCAA 840
 |||||||
 Db 968 GATTACACACAGTGAACACCTTCTAAGTTTGATTAACGGCAATTTTATTTCTCTCAA 1027
 |||||||
 Oy 841 CGACGCGCAGAGTGTCCCGCAGCGCGCCACCCCGTTATTTGATTTGATCAGTCAGCG 900
 |||||||
 Db 1028 CGACGCGCAGAGTGTCCCGCAGCGCGCCACCCCGTTATTTGATTTGATCAGTCAGCG 1087
 |||||||
 Oy 901 TTGACGCGCCATCCACCGCAAAAACAGGCGGTATGATGATTAACCATCTTCAGTACG 960
 |||||||
 Db 1088 TTGACGCGCCATCCACCGCAAAAACAGGCGGTATGATGATTAACCATCTTCAGTACG 1147
 |||||||
 Oy 961 TTTATTTGCGGACAGATTAATCTTGCAATATCGCGGCGGCTGAGACTGCAACTG 1020
 |||||||
 Db 1148 TTTATTTGCGGACAGATTAATCTTGCAATATCGCGGCGGCTGAGACTGCAACTG 1207
 |||||||
 Oy 1021 ACGCTTCCCGGTGACCGGATTAACAGCGCGCAGGTGATGATGATGTTGAACGCTGCG 1080
 |||||||
 Db 1208 ACGCTTCCCGGTGACCGGATTAACAGCGCGCAGGTGATGATGATGTTGAACGCTGCG 1267
 |||||||

QY 1081 CGTCGGCTAAGCATACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1140
 |||||||
 Db 1268 CGTCGGCTAAGCATACAGCCAGTGGATTTCAGGTTTCGCTGGTCTTCCAGACTTTACAG 1327
 |||||||
 QY 1141 CAGATGCGTGAATAAACGCCGCTGTCATTAAATACGCCGCCCGAGAGGTGAACACTGACC 1200
 |||||||
 Db 1328 CAGATGCGTGAATAAACGCCGCTGTCATTAAATACGCCGCCCGAGAGGTGAACACTGACC 1387
 |||||||
 QY 1201 CTGGCAGATGTAAGAGCGAATGGCGAGGCGATGTGTTGTTGGCAGGTTTACGCAA 1260
 |||||||
 Db 1388 CTGGCAGATGTAAGAGCGAATGGCGAGGCGATGTGTTGTTGGCAGGTTTACGCAA 1447
 |||||||
 QY 1261 ATCGTAATGAAGACAGCATACCGGCGTGCAGTTGAGAT 1300
 |||||||
 Db 1448 ATCGTAATGAAGACAGCATACCGGCGTGCAGTTGTAAT 1487
 |||||||

Search completed: June 12, 2003, 11:33:25
 Job time : 3573 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 09:19:38 ; Search time 2118 Seconds
(Without alignments)
10116.444 Million cell updates/sec

Title: US-09-777-566a-1
Perfect score: 1323
Sequence: 1 atgaagcgccttaaccc.....atcacatcacatcactaa 1323

Scoring table: IDENTITY_NUC
Gapop.10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614.4	46.4	616	BJ074127	BJ074127 BJ074127
2	386.2	29.2	393	BE520240	BE520240 M1B12STM
3	383	28.9	383	AM036132	AM036132 EST274508
4	383	28.9	383	AM036134	AM036134 EST274510
5	304.6	23.0	354	AA545747	AA545747 HBMSF1B4-
6	216.8	16.4	320	BE436403	BE436403 EST407481

Result No.	Score	Query Match	Length	DB ID	Description
7	208	15.7	208	BM409016	BM409016 EST583343
8	208	15.7	208	BM412806	BM412806 EST587133
9	198.8	15.0	546	BJ040796	BJ040796 BJ040796
10	169.4	12.8	179	BE461872	BE461872 EST413291
11	160.8	11.9	508	BJ030203	BJ030203 BJ030203
12	157.6	11.2	657	BM457132	BM457132 NF10C05P
13	143.2	10.8	352	BE520241	BE520241 M1B12XTM
14	43.8	3.3	394	BE520241	BE520241 M1B12XTM
15	41.2	3.1	733	BM60613	BM60613 YF44G12.s1
16	39.8	3.0	431	BM152280	BM152280 602917706
17	39.8	3.0	431	AT553154	AT553154 V107B12.y
18	39.8	3.0	602	BB610221	BB610221 BB610221
19	39.8	3.0	607	AA521521	AA521521 V107B12.y
20	39.2	3.0	553	AM390601	AM390601 MR3-ST019
21	38.6	2.9	398	AA15940	AA15940 on18d04.s
22	38.4	2.9	669	AA596678	AA596678 1M0410006
23	38.2	2.9	379	BE289814	BE289814 601089743
24	38.2	2.9	499	AT930492	AT930492 U165e01.y
25	38.2	2.9	693	BM944391	BM944391 U1-M-EHOP
26	38.2	2.9	693	BM950172	BM950172 U1-M-EHOP
27	38.2	2.9	694	BM611178	BM611178 BM611178
28	38.2	2.9	756	BM945032	BM945032 U1-M-EHOP
29	38.2	2.9	900	BM951509	BM951509 AGENCOURT
30	37.8	2.9	942	BM1653984	BM1653984 603280045
31	37.6	2.8	274	BM153360	BM153360 NF035C07I
32	37.6	2.8	553	BM134827	BM134827 U1-H-B11-
33	36.8	2.8	618	BM938287	BM938287 mgyb00196
34	36.4	2.8	875	BM182891	BM182891 601809824
35	36.2	2.7	674	BM93293	BM93293 CIT-HSP-217
36	36.2	2.7	764	BM536305	BM536305 602051954
37	36	2.7	638	BM486941	BM486941 pgm2n.pk0
38	36	2.7	790	BM4293058	BM4293058 BJ_Ba000
39	35.8	2.7	957	BM581744	BM581744 602099525
40	35.6	2.7	396	AA479432	AA479432 zv14d11.r
41	35.4	2.7	238	AA615254	AA615254 h427h06.x
42	35.4	2.7	295	AA318969	AA318969 EST21203
43	35.4	2.7	389	AA669772	AA669772 ad41h08.s
44	35.4	2.7	475	AA243772	AA243772 z67d01.r
45	35.4	2.7	483	AM609430	AM609430 MR3-ST020

ALIGNMENTS

RESULT 1
BJ074127/c 616 bp mRNA linear EST 11-DEC-2001
LOCUS BJ074127 NIBB Mochil normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone XJ090j06 5', mRNA sequence.
ACCESSION BJ074127
VERSION BJ074127.1 GI:17504316
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yatai, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source
Location/Qualifiers
1..616
/organism="Xenopus laevis"
/db_xref="taxon:8355"

/clone="X1090j06"
 /clone_lib="NIBB Mochil normalized Xenopus tailbud
 library"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAs
 were oligo-dt primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subcloned
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."

BASE COUNT 128 a 167 c 167 g 154 t
 ORIGIN

Query Match 46.4%; Score 614.4; DB 13; Length 616;
 Best Local Similarity 99.8%; Pred. No. 5.1e-190;
 Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GGTAGCCGAGGATGCTGGCGAAAAAGGCTCCCGCACTGTCAGTCCGCAATT 323
 DB 616 GGTAGCCGAGGATGCTGGCGAAAAAGGCTCCCGCACTGTCAGTCCGCAATT 557
 QY 324 TGCTGATGTGACGACGATGCTACCCGTAACAGGCGAAGCTTCGCGCGCTGGCACC 383
 DB 556 TGCTGATGTGACGACGATGCTACCCGTAACAGGCGAAGCTTCGCGCGCTGGCACC 497
 QY 384 TGAAGTGCATTAACGCTACATACCCAGGAGATAGCTCCAGTCCGATCCGTTATTAA 443
 DB 496 TGAAGTGCATTAACGCTACATACCCAGGAGATAGCTCCAGTCCGATCCGTTATTAA 437
 QY 444 TCCTCTAAACAGTGGGCTTGGCAACTGATGATACGCGAGATGATGATGATGATGATGAT 503
 DB 436 TCCTCTAAACAGTGGGCTTGGCAACTGATGATGATGATGATGATGATGATGATGATGAT 377
 QY 504 CAGGCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563
 DB 376 CAGGCGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 317
 QY 564 GGAAGCGGCTTAATTTTCCGCAATCAACTTGTGCTTAAAGCTGAGAAAGAGAGA 623
 DB 316 GGAAGCGGCTTAATTTTCCGCAATCAACTTGTGCTTAAAGCTGAGAAAGAGAGA 257
 QY 624 AAGCTGTCATTAACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 683
 DB 256 AAGCTGTCATTAACGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 197
 QY 684 ATTAACCGGTGCGGTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 743
 DB 196 ATTAACCGGTGCGGTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 137
 QY 744 ACAGGAGATGCGGAGCGGGGTGGGAGAGATCACCAGATTACACCAAGTGAACACCTT 803
 DB 136 ACTGGGAATGCGGAGCGGGGTGGGAGAGATCACCAGATTACACCAAGTGAACACCTT 77
 QY 804 GCTAAGTTGCTAATGAGGCAATTTTATTTGCTCAACAGCAGCAGAGAGTGGCCGAG 863
 DB 76 GCTAAGTTGCTAATGAGGCAATTTTATTTGCTCAACAGCAGCAGAGAGTGGCCGAG 17
 QY 864 CCGCGCAACCCCGTTA 879
 DB 16 CCGCGCAACCCCGTTA 1

RESULT 2
 BE520240/c 393 bp mRNA linear EST 19-MAR-2001
 LOCUS M1B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
 DEFINITION clone M1B12 5', mRNA sequence.
 ACCESSION BE520240
 VERSION BE520240.1 GI:9778242
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Contact: Benning, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA. FAX:
 6142920603 TEL: 6142929371.

FEATURES

Location/Qualifiers
 1..393
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_lib="M1B12"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"

BASE COUNT

86 a 100 c 103 g 102 t 2 others

Query Match 29.2%; Score 386.2; DB 10; Length 393;
 Best Local Similarity 98.7%; Pred. No. 3.5e-115;
 Matches 388; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 319 ATTAATTCCTGATGTCGACGAGCTACCCGTAACACGCGCAAGCTTCGCGCGGCTG 378
 DB 393 ATTAATTCCTGATGTCGACGAGCTACCCGTAACACGCGCAAGCTTCGCGCGGCTG 334
 QY 379 GCACCTGACGTGCAATTAACGCTACATACCCAGGAGATAGCTCAGTCCGATCCGTTA 438
 DB 333 GCACCTGACGTGCAATTAACGCTACATACCCAGGAGATAGCTCAGTCCGATCCGTTA 274
 QY 439 TTTATTCCTCTAAACATGCGGCTTGGCACTGATTAACGGAACGTGACTGACGCGATC 498
 DB 273 TTTATTCCTCTAAACATGCGGCTTGGCACTGATTAACGGAACGTGACTGACGCGATC 214
 QY 499 CTCAGCAGGCGAGAGGCTCAATGCTGACTTACCGGCGCATGCGCAAGCGGTTGCG 558
 DB 213 CTCAGCAGGCGAGAGGCTCAATGCTGACTTACCGGCGCATGCGCAAGCGGTTGCG 154
 QY 559 GAAGTGAAGCGGCTTAATTTTCCGCAATCAACTGTGCTTAAACGTGAGAAACAG 618
 DB 153 GAAGTGAAGCGGCTTAATTTTCCGCAATCAACTGTGCTTAAACGTGAGAAACAG 94
 QY 619 GACGAAACCTGTTATTAACGAGGATTAACATGGAATCAAGGTGAGCGGACAAAT 678
 DB 93 GACGAAACCTGTTATTAACGAGGATTAACATGGAATCAAGGTGAGCGGACAAAT 34
 QY 679 GTCTCATTAACCGGTGCGGTAAAGCTTCGCATCA 711
 DB 33 GTCTCATTAACCGGTGCGGTAAAGCTTCGCATCA 1

RESULT 3
 AW036132 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AW036132
 DEFINITION EST74508 tomato seed, TAMD lycopersicon esculentum cDNA clone

cLEE1E23 similar to periplasmic phosphoanhydride phosphohydrolase precursor, putative, mRNA sequence.
 AM036132
 AM036132.1 GI:5894811
 EST.
 SOURCE
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)
 Alcala, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E., Liang, F., Upton, J., Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 CONTACT: CUGI
 CLEMSON UNIVERSITY GENOMICS INSTITUTE
 100 JORDAN HALL, CLEMSON, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 3 prime sequence.
 Location/Qualifiers
 1..383
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEE1E23"
 /clone_1id="tomato seed, TAMU"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="X1L-Blue MR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEE - Tomato Seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."
 EcoRI 101 c 97 g 83 t
 BASE COUNT
 ORIGIN
 Query Match 28.9%; Score 383; DB 10; Length 383;
 Best Local Similarity 100.0%; Pred. No. 3.9e-114;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 348 TAAACGAGGGAAGCTTCCGCGCGGCTGGACCTGACTGTGCATTAACCGTACATC 407
 DB 1 TAAACGAGGGAAGCTTCCGCGCGGCTGGACCTGACTGTGCATTAACCGTACATC 60
 QY 408 CCAAGGAGATACGTCGATCCGCGGCTGTTATTTAATCCTTAATAAAGTGGCTTGGCA 467
 DB 61 CCAAGGAGATACGTCGATCCGCGGCTGTTATTTAATCCTTAATAAAGTGGCTTGGCA 120
 QY 468 ACTGATTAACGCGAAGCTGACTGACGCGATCCTTCAGCAGGAGGAGGTCATTTGCTGA 527
 DB 121 ACTGATTAACGCGAAGCTGACTGACGCGATCCTTCAGCAGGAGGAGGTCATTTGCTGA 180
 QY 528 CTTTACCGGCGATCGGCAAAAGCGGCTTCCGGAACGTGAACGGGTGCTTAATTTCCGCA 587
 DB 181 CTTTACCGGCGATCGGCAAAAGCGGCTTCCGGAACGTGAACGGGTGCTTAATTTCCGCA 240
 QY 588 ATCAAACTTGTCCTTAAGCTGAGAAACGAGGAAAGCTGTCATTAACGAGGATC 647
 DB 241 ATCAAACTTGTCCTTAAGCTGAGAAACGAGGAAAGCTGTCATTAACGAGGATC 300
 QY 648 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGTGGGTAAACCTTCGC 707
 DB 301 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGTGGGTAAACCTTCGC 360
 QY 708 ATCAATGCTGAGGAGATATTTTC 730
 DB 361 ATCAATGCTGAGGAGATATTTTC 383

RESULT 4
 AM036134/c
 LOCUS
 DEFINITION
 EST274510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
 cLEE1E23 similar to periplasmic phosphoanhydride phosphohydrolase precursor, putative, mRNA sequence.
 AM036134
 AM036134.1 GI:5894813
 EST.
 SOURCE
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 383)
 Alcala, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E., Liang, F., Upton, J., Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 CONTACT: CUGI
 CLEMSON UNIVERSITY GENOMICS INSTITUTE
 100 JORDAN HALL, CLEMSON, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 Location/Qualifiers
 1..383
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cLEE1E23"
 /clone_1id="tomato seed, TAMU"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="X1L-Blue MR"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEE - Tomato Seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."
 EcoRI 83 a 97 c 101 g 102 t
 BASE COUNT
 ORIGIN
 Query Match 28.9%; Score 383; DB 10; Length 383;
 Best Local Similarity 100.0%; Pred. No. 3.9e-114;
 Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 348 TAAACGAGGGAAGCTTCCGCGCGGCTGGACCTGACTGTGCATTAACCGTACATC 407
 DB 383 TAAACGAGGGAAGCTTCCGCGCGGCTGGACCTGACTGTGCATTAACCGTACATC 324
 QY 408 CCAAGGAGATACGTCGATCCGCGGCTGTTATTTAATCCTTAATAAAGTGGCTTGGCA 467
 DB 323 CCAAGGAGATACGTCGATCCGCGGCTGTTATTTAATCCTTAATAAAGTGGCTTGGCA 264
 QY 468 ACTGATTAACGCGAAGCTGACTGACGCGATCCTTCAGCAGGAGGAGGTCATTTGCTGA 527
 DB 263 ACTGATTAACGCGAAGCTGACTGACGCGATCCTTCAGCAGGAGGAGGTCATTTGCTGA 204
 QY 528 CTTTACCGGCGATCGGCAAAAGCGGCTTCCGGAACGTGAACGGGTGCTTAATTTCCGCA 587
 DB 203 CTTTACCGGCGATCGGCAAAAGCGGCTTCCGGAACGTGAACGGGTGCTTAATTTCCGCA 144
 QY 588 ATCAAACTTGTCCTTAAGCTGAGAAACGAGGAAAGCTGTCATTAACGAGGATC 647
 DB 143 ATCAAACTTGTCCTTAAGCTGAGAAACGAGGAAAGCTGTCATTAACGAGGATC 84
 QY 648 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGTGGGTAAACCTTCGC 707
 DB 83 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGTGGGTAAACCTTCGC 24

OY 708 ATCAATGCTGACGAGATATTC 730
 DB 23 ATCAATGCTGACGAGATATTC 1

RESULT 5
 AA545747/c 354 bp mRNA 11linear EST 12-MAY-1999
 LOCUS HBMSFLB4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens CDNA
 DEFINITION clone HBMSFLB4 5', mRNA sequence.

ACCESSION AA545747
 VERSION AA545747.1 GI:2307026
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 354)
 Jia, L., Robey, P., Young, M., Schuler, G., Powell, J., Yang, L., Lennon
 G., Hillier, L., Allen, M., Bowles, L., Gelsel, S., Kucaba, T., Marra
 M., Martin, J., Steptoe, M., Tan, F., Theising, B., Bowers, Y., Wylie
 T., Waterston, R., Wilson, R. and Francomano, C.
 WashU-MGB/NHGRI EST Project
 UNPUBLISHED (1997)
 CONTACT: Ldbin Jia
 Medical Genetics Branch
 National Human Genome Research Institute
 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
 Tel: 301-402-4877
 Fax: 301-496-7157
 Email: lbinhelix.nih.gov
 Seq primer: M13 Reverse.

FEATURES
 source
 1..354
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HBMSFLB4"
 /clone_1id="Human Bone Marrow Stromal Fibroblast"
 /sex="Male and Female"
 /tissue_type="bone marrow"
 /cell_type="stromal fibroblast"
 /dev_stage="mixed"
 /lab_host="XLI-Blue"
 /note="Vector: pBluescript, Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 77 a 95 c 95 t 2 others

ORIGIN

Query Match 23.0% Score 304.6; DB 9; Length 354;
 Best Local Similarity 94.9%; Pred. No. 2.1e-88;
 Matches 333; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

OY 497 TCCTCAGCAGGCGAGAGGCTCAATTCCTGACTTACCGGGCATCGCAAGCGCGTTTC 556
 DB 351 TCTTCACAGGCGAGGCGA-GGTCAATTCCTGACTTACCGGGCATCGCAAGCGCGTTTC 293

OY 557 GCGAAGTGGAGCGGCTTAATTTCCGCAATCAACTTGGCTTAAGAGTGAAGAC 616
 DB 292 GCGAAGTGGAGCGGCTG-TTAATTTCCGCAATCAACTTGGCTTAAGAGTGAAGAC 234

OY 617 AGGAGGAAGGCTGCTTACGAGGATTAACCATGGAAGTGAAGGAGCGGCA 676
 DB 233 AGGAGGAAGGCTGCTTACGAGGATTAACCATGGAAGTGAAGGAGCGGCA 174

OY 677 ATGTCATTAACCGGTGCGGTAGCCCTCGCATCATGCTGACGAGATATTTCTCTCG 736
 DB 173 ATGTCATTAACCGGTGCGGTAGCCCTCGCATCATGCTGACGAGATATTTCTCTCG 114

OY 737 AACAGACAGAGGATGCGGAGCGGCGGTGGGAGAGATACCGATTACACCACTGGA 796
 DB 113 AACAGACAGAGGATGCGGAGCGGCGGTGGGAGAGATACCGATTACACCACTGGA 54

OY 797 ACACCTGCTAAGTTTGATACAGCGCAATTTATTCTCAACGACGCGCA 849

DB 53 ACACCTGCTAAGTTTGATACAGCGCAATTTATTGCTACACGACGCGCA 1

RESULT 6
 BE436403/c 320 bp mRNA 11linear EST 18-MAY-2001
 LOCUS EST407481 tomato breaker fruit, TIGR Lycopersicon esculentum CDNA
 DEFINITION clone CLEG32C20, mRNA sequence.

ACCESSION BE436403
 VERSION BE436403.1 GI:9434246
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE
 1 (bases 1 to 320)
 Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
 Jiang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Romling, C.M.,
 Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
 S.D.
 Generation of ESTs from tomato fruit tissue, breaker stage
 UNPUBLISHED (2000)
 CONTACT: CUCI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source
 1..320
 /organism="Lycopersicon esculentum"
 /cultivar="T49496"
 /db_xref="taxon:4081"
 /clone="CLEG32C20"
 /clone_1id="tomato breaker fruit, TIGR"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCudaapt; Site_1: EcoRI;
 Site_2: XhoI; Fruit were harvested at the breaker stage
 (first sign of lycopene accumulation on the blossom end
 of the fruit). Fruit were cut in half and the seeds and
 locules were discarded prior to freezing the pericarp."

BASE COUNT 63 a 96 c 86 g 75 t

ORIGIN

Query Match 16.4% Score 216.8; DB 10; Length 320;
 Best Local Similarity 95.1%; Pred. No. 1.4e-59;
 Matches 289; Conservative 0; Mismatches 7; Indels 8; Gaps 6;

OY 242 GACATTAACAGCGGCGTGTGATGCGGAGATGCTGGCAAAAGGCGTCCCGC 301
 DB 320 GACATTAACAGCGGCGTGTGATGCGGAGATGCTGGCAAAAGGCGTCCCGC 261

OY 302 AGTCTGTCAGTGGCGGATTAATTCCTGATGTCGAGAGCG--TACCCGTAAGAGCGGA 359
 DB 260 AGTCTGTCAGTGGCGGATTAATTCCTGATGTCGAGAGCGGTTACCCGTAAGAGCGGA 201

OY 360 AGCCTGCGCCCGCGGCGTGGCAGCTGATGCAATAACCGTACATACCAAGCAGATAC 419
 DB 200 AGCCTGCGCCCGCGGCGTGGCAGCTGATGCAATAACCGTACATACCAAGCAGATAC 141

OY 420 G--TCCAGTCCCGATCGTTATTTAA--TCCCTTAATAAACTGGC-GTTTGCCAACTGATA 475
 DB 140 GTCACAGGTCCCGATCGTTATTTAAATTCCTTAATAAACTGGCCTTGGCCAACCTGATA 81

OY 476 ACAGGAGAGTACAGAGCGGA--TCCTCAGCAGGGGAGG--AGGGTCAATTTGCTGACTTAC 533
 DB 80 ACAGGAGAGTACAGAGCGGATTCCTTCAGCAGGGGAGGAGGTCATTTGCTGACTTAC 21

QY 534 CGGG 537
 Db 20 CGGG 17

RESULT 7
 LOCUS BM409016/c
 DEFINITION EST583343 tomato breaker fruit Lycopersicon esculentum cDNA clone
 ACCESSION CLE66H24 5' end, mRNA sequence.
 VERSION BM409016
 KEYWORDS GI:18260646
 SOURCE EST.
 ORGANISM Lycopersicon esculentum
 tomato.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 208)

REFERENCE
 AUTHORS Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai
 J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning
 C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 JOURNAL Unpublished (2002)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3

FEATURES
 source
 Location/Qualifiers

1..208
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLE66H24"
 /clone_lib="tomato breaker fruit"
 /tissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCunadap1; Site.1: EcoRI;
 Site.2: XhoI; supplier: Boyce Thompson Institute; Fruit
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 lycopene accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."
 BASE COUNT 43 a 56 c 62 g 47 t
 ORIGIN

Query Match 15.7%; Score 208; DB 13; Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4e-57;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 301
 |||||||
 Db 208 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 149
 QY 302 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 361
 |||||||
 Db 148 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 89
 QY 362 CCTTCCGCGCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCCAGCAGATACGT 421
 |||||||
 Db 88 CCTTCCGCGCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCCAGCAGATACGT 29
 QY 422 CCAGTCCGATCCGTTATTTAATCCTCT 449
 |||||||
 Db 28 CCAGTCCGATCCGTTATTTAATCCTCT 1

RESULT 8
 LOCUS BM412806/c
 DEFINITION EST587133 tomato breaker fruit Lycopersicon esculentum cDNA clone
 ACCESSION CLE66I013 5' end, mRNA sequence.
 VERSION BM412806
 KEYWORDS GI:18264436
 SOURCE EST.
 ORGANISM Lycopersicon esculentum
 tomato.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 208)

REFERENCE
 AUTHORS Alcalá, J., Vrebalov, J., White, R., Vision, T., Karamycheva, S.A., Tsai
 J., Bougri, O., Kirkness, E., Utterback, T., Van Aken, S., Ronning
 C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 JOURNAL Unpublished (2002)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3.

FEATURES
 source
 Location/Qualifiers

1..208
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLE66I013"
 /clone_lib="tomato breaker fruit"
 /tissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCunadap1; Site.1: EcoRI;
 Site.2: XhoI; supplier: Boyce Thompson Institute; Fruit
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 lycopene accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."
 BASE COUNT 43 a 56 c 62 g 47 t
 ORIGIN

Query Match 15.7%; Score 208; DB 13; Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4e-57;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 301
 |||||||
 Db 208 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 149
 QY 302 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 361
 |||||||
 Db 148 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 89
 QY 362 CCTTCCGCGCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCCAGCAGATACGT 421
 |||||||
 Db 88 CCTTCCGCGCGGGGCTGGACCTGACCTGTGCAATTAACCGTACATACCCAGCAGATACGT 29
 QY 422 CCAGTCCGATCCGTTATTTAATCCTCT 449
 |||||||
 Db 28 CCAGTCCGATCCGTTATTTAATCCTCT 1
 RESULT 9
 LOCUS BJ040796
 BJ040796 546 bp mRNA linear EST 06-DEC-2001

REFERENCE AUTHORS	TITLE JOURNAL COMMENT	FEATURES source	BASE COUNT ORIGIN
Lycopodium. 1 (bases 1 to 179) Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Tiang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romling,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley S.D.	Generation of ESTs from tomato fruit tissue, breaker stage Unpublished (2000) Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence: Location/Qualifiers	1. /179 /organism="Lycopodium esculentum" /cultivar="T496" /db_xref="taxon:4081" /clone="CLEG40G22" /clone_lib="tomato breaker fruit, TIGR" /tissue_type="Pericarp" /tissue="breaker" /dev_stage="SOLR" /lab_host="SOLR" /note="Vector: BluescriptSKmCnadapt; Site.1: SCor1; Site.2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."	40 a 47 c 51 g 41 t
Query Match	12.8%; Score 169.4; DB 10; Length 179;		
Best Local Similarity	96.6%; Pred. No. 3.7e-44;		
Matches 173; Conservative	0; Mismatches 6; Indels 0; Gaps 0;		
OY	398 CCGTACATACCAGGACAGATAGTCCTCCGATCCGTATTATATCTCTTAAAACTG 457		
DB	179 CCGTACATACCAGGACAGATAGTCCTCCGATCCGTATTATATCTCTTAAAACTG 120		
OY	458 GCGTTTGCCAACTGGATTAACGCGAAGCTGATCGACGATCCTCGACGAGGAGGAGGT 517		
DB	119 GCGTTTGCCAACTGGATTAACGCGAAGCTGATCGACGATCCTCGACGAGGAGGAGGT 60		
OY	518 CAATTGCTGACTTATCCGGCATCGGCAAGGCGCTTTCCGAACTGGAACGGGTCTT 576		
DB	59 CAATTGCTGACTTATCCGGCATCGGCAAGGCGCTTTCCGAACTGGAACGGGTCTT 1		
RESULT 11			
BJ030203	508 bp mRNA linear EST 05-DEC-2001		
LOCUS	BJ030203 N18B Mochl normalized Xenopus neurola library Xenopus		
DEFINITION	laevis cDNA clone XL003H05 5', mRNA sequence.		
ACCESSION	BJ030203		
VERSION	BJ030203.1		
KEYWORDS	GI:17371629		
SOURCE	EST.		
ORGANISM	African clawed frog.		
	Xenopus laevis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
REFERENCE	Xenopodinae; Xenopus.		
AUTHORS	1 (bases 1 to 508)		
	Kitayama,A., Terasaka,C., Mochi,N., Ueno,N., Shin-I,T. and Kohara Y.		
TITLE	Expressed genes in X. laevis embryo		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel.: 81-559-81-6856 Fax.: 81-559-81-6855		

[illegible]

Db	131	CGCCCCACATGATGCTGTGCGGAGTCAACCGCGGCACTGATCATCTGCGCTGTG	72
QY	244	CATTACCAAGCGAGGCTGTGTAAGCCGAGGATTTGCTG	282
Db	71	AGGTCCGACGTGAGCGGCGCTTCGAGTTCAACAGAGCTG	33

Search completed: June 12, 2003, 12:09:00
 Job time : 2128 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 13:01:29 ; Search time 3587 Seconds

(Without alignments)
3569.901 Million cell updates/sec

Title: US-09-777-566A-2

Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLPQSA.....IVNEARIPACSLRHHHHH 440

Scoring table: BIOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-rlh
-Q/cgna2_1/USFRO.spool/US09777566/runat_05062003_154808_6550/app.query.fasta.1.583
-DB-GenDbml -QWt-fastap -SUFFIX-rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-Dlosum62 -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09777566 @CGN 1.1 2496 @runat_05062003_154808_6550 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	6	AR108133 Sequence
2	2302	100.0	1323	6	AR127818 Sequence
3	2302	100.0	1323	6	AR130956 Sequence
4	2302	100.0	1323	6	AX052713 Sequence
5	2302	100.0	1323	6	AX338966 Sequence
6	2302	100.0	1323	6	AX356566 Sequence
7	2235	97.1	1901	1	ECOAPPPA
8	2235	97.1	1901	6	AX356572
9	2235	97.1	3470	6	AX042376
10	2235	97.1	4060	6	AX042374
11	2235	97.1	6116	6	AX042375
12	2235	97.1	6708	6	AX042373
13	2235	97.1	10784	6	AE000200
14	2235	97.1	17732	6	AX042378
15	2235	97.1	18882	1	D90735
16	2235	97.1	20623	6	AX042372
17	2229	96.8	1901	1	ECOACPHA
18	2229	96.8	1901	1	ECOACPHC
19	2229	96.8	1901	1	ECOACPHD
20	2229	96.8	1901	1	ECOACPHB
21	2226	96.7	1901	6	AX356574
22	2226	96.7	1901	1	ECOACPHF
23	2221	96.5	1901	6	AX356571
24	2221	96.5	1901	6	AX356570
25	2210	96.4	1901	6	AE005292
26	2210	96.0	10029	1	AE002098
27	2210	96.0	327773	1	AP002554
28	2205	95.8	5421	6	AX042377
29	1102.5	47.9	212936	2	AC020970
30	943.5	41.0	10029	1	AE013783
31	943.5	41.0	193050	1	AJ414149
32	570	24.8	114979	2	AC020798
33	545	23.7	10278	1	AE005294
34	541	23.5	1810	1	PRAGRP
35	539	23.4	1675	1	ECOAGPA
36	539	23.4	11710	1	AE000202
37	539	23.4	18252	1	D90737
38	529	23.0	21670	1	AE008748
39	527	22.9	700	1	ECAPPA
40	526	22.8	10545	1	AE005727
41	524	22.8	234050	1	AL627269
42	505	21.9	10385	1	AE011713
43	505	21.9	13498	1	AE012175
44	483.5	21.0	7129	1	AE012063
45	373	16.2	11877	1	AE011683

RESULT 1

ALIGNMENTS

AE011683 Xanthomon

LOCUS	AR108133	1323 bp	DNA	linear	PAT 14-FEB-2001
DEFINITION	Sequence 1 from patent US 6110719.				
ACCESSION	AR108133				
VERSION	AR108133.1	GI:12823620			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1323)				
AUTHORS	Kretz, K.				
TITLE	Phytase				
JOURNAL	Patent: US 6110719-A 1 29-AUG-2000;				
FEATURES	location/qualifiers				
source	1..1323				
BASE COUNT	323 a 353 c 357 g 289 t			1 others.	
ORIGIN	/organism="unknown"				
Alignment Scores:					
Pred. No.:	3.27e-169	Length:	1323		
Score:	2302.00	Matches:	440		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-777-566a-2 (1-440) x AR108133 (1-1323)					
QY	1 MethysalalieleuileProPheleuserleuileProleuThrProGlnSera	20			
Db	1 ATGAAGGAGTCTAATCCATTTTATCTCTTCGATCCGTAAACCCGCAATCTGA	60			
QY	21 PhealaginsergluProGluLeuLysleuGlnSerValValIleValSerArgHisGly	40			
Db	61 TTCCTCGAGATGAGCCCGGACCTGAAAGTGGTGATGTGATGTCATGTCATGGT	120			
QY	41 ValArgAlaProThrIlePheGlnleuMetGlnAspValThrProAspAlaTrpPro	60			
Db	121 GTGGCTGCTCCAAACCAAGCCACGCACTGATGAGATGTCACCCAGACCAATGGCCA	180			
QY	61 ThrTPProValLysleuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpIleu	80			
Db	181 ACCTGGCCGGTAAACCTGGTGGCTGACACCCGCGNGTGGTGGATCACTATCCATCTC	240			
QY	81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyGlyCysPro	100			
Db	241 GGACATTAACCAACGCCACGCTGCTGTACCCACGATGCTGGCGAAAGGCGTGC	300			
QY	101 GlnserGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu	120			
Db	301 CAGCTGTGTCAGGTCGCCATTAATGCTGATGTCGACGAGCGGACCCGTAACACAGCGCA	360			
QY	121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140			
Db	361 GCCTTCGGCCGGGGCTGGCACCTGACGTGCATTAACACGTAACACCAAGCAATACG	420			
QY	141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160			
Db	421 TCCAGTCCCGGTCGGTATTATTAATCCCTTAAAACTGGCGTTGCCCACTGGATTAACCG	480			
QY	161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180			
Db	481 AACGTGATGACGGGATCTCAGCACGGGACAGAGGGTCAATTTGCTGACTTAAACCGGCAT	540			
QY	181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly	200			
Db	541 CGGCAACGCGGCTTCGGGAACCTGGAACGGGTCTTAATTTTCCCAATCAAACTTGTGC	600			
QY	201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220			
Db	601 CTTAAAGCTGAGAAACAGACGAAAGCTGTTCATTAAACGACGAGGATTAACATCGGAAC	660			

QY	221	lyslalserlalaaspaasvalserleuthrclylavalserleualasermetleuthr	240
Db	661	AAGGTGACGGCCGACAAATGCTCAATTAACCGGTCGGGTAAGCCTTCGACATCATCTGACG	720
QY	241	GlulIephelauLeuGInGlnAlaGInglyMetProGluProGlyTTPGlyArgIlethr	260
Db	721	GAGMTATTTTCCTCGACAAACAGACAGGAGTGGCGGAGCCGGGGTGGGAAAGATCACC	780
QY	261	AspSerHisGlnTTPasrNthrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
Db	781	GATTCACACACAGTGGACACACCTTCCTAAGTTTGCATTAACGGCAGATTTTATTTCTCTCAA	840
QY	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	841	CGCAGCGCAGAGGTTGCCCGACGGCCGCCACCCCGTTATTTGGATTGATATATGGCAGCG	900
QY	301	LeuThrProHisProProGlnIlySGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
Db	901	TTTGACGCCCATCCACCCGAAAACAGCGCATGTGCTGGACATTAACCACTTCAGTACTG	960
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTTP	340
Db	961	TTTATTGCGCGACACGATACTAAATCTGGCAAAATCTCGCGCGCGACATGAGACTCACTGG	1020
QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTTP	360
Db	1021	ACGCTTCCCGGTACCGCGGATTAACCGCCGCGACGGTGTGAATGtggttTGAACGGCTGG	1080
QY	361	ArgArgLeuSerAspAsnSerGlnTTPIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	1081	CGTGGGGTAAACGATTAACAGCCAGTGGATTAGGTTTCGCTGGCTTCACAGACTTTACAG	1140
QY	381	GlnMetArgAspIlysrThrProLeuSerLeuAsnThrProProGlyGluValIlyLeuThr	400
Db	1141	CAGATGCGGTGTAATAACCCCGCTGTCAATTAATACGCCGCCGCGGAGAGGTGMAACTGACC	1200
QY	401	LeuAlaGlyCysGluGluArgAsnAlaGlnIlyMetCysSerLeuAlaGlyPheThrGln	420
Db	1201	CTGGCAGGATGTGAAGACCGAAATATGCCAGAGGCATGTGTTCTGTTGGCAGGTTTATACGCA	1260
QY	421	IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis	440
Db	1261	ATCGTGATGAAGACCGCATACCGCGGTGAGTTTGATGATCTCATCAACATCACCATTCAC	1320
RESULT 2			
ARI27818			
LOCUS	ARI27818	1323 bp	DNA
DEFINITION	Sequence 1 from patent US 6183740.		linear
ACCESSION	ARI27818		
VERSION	ARI27818.1	GI:14115480	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1323)		
TITLE	Short,J.M. and Kretz,K.A.		
JOURNAL	Recombinant bacterial phytases and uses thereof		
FEATURES	Patent: US 6183740-A 1 06-FEB-2001;		
source	Location/Qualifiers		
	1..1323		
	/organism="unknown"		
BASE COUNT	323 a 353 c 357 g	289 t	1 others
ORIGIN			
Alignment Scores:			
Pred. No.:	3,27e-169	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

QY 1 MettysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGAGCATTTATCCATTTTATCTCTTCGATTCCTTAACCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGCTGTGATGTCAGTGTGATGCT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 121 GTGCTGTCTCCAAACCAAGGCCAGCACTGATGAGGATGTCAACCCAGACCAATGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLysLeuIleAlaTrpLeu 80
 DB 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCCGNGTGGTGGAGCTAATGCCATATCTC 240
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 241 GGACATTTACCAACGCCAGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGGCTGCCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTGTGTGACAGTGGCATTATTCCTGATGTGACAGACCGTACCCGTAACACAGCGCA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCCCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGCCAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCACTCCCGATCCGTTATTTATCTCTAAACCTGGGGTGTGGCACTGTGATTAACGG 480
 QY 161 AsnValIleThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACGTACGCGCATCTCAGCAGGCGAGAGGCTCAATGTGACTTTCGCGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAAAAGCGGCTTTCGGGACTGGAACGGGTGCTTAATTTTCCGCAATCAACTGTGTC 600
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTACGACGACATTAACCATCGGAATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 QY 241 GlnIlePheLeuLeuGlnGlnIleAlaGlnGlyMetProGluProGlyTrpProLysArgIleThr 260
 DB 721 CAGATATTTCTCTCAACACACAGGAAATGCCGAGGCCGGGGGGGGAAGGATCACCC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 781 GATTCACACACAGTGAACACCTTGTGTAAGTTGCAATACGGCAATTTTATTTGGTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCGCAGAGGTGGCCCGCAGCGCGCACCCGCTTAATGATTTGATTCATGTCACAGC 900
 QY 301 LeuThrProHisProProGlnLysGlnIleAlaTrpGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATTCACCGCAAAACAGGGGTATGGTGTGACATTTACCCACTTCAGTACACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuLeuAsnTrp 340
 DB 961 TTATATGGCGGACAGATCTAATCTGGCAAAATCTCGGGGCGCACTGGAAGTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgThr 360
 DB 1021 ACGCTTCCCGGTACACCGCATTAACAGCGCGCAGGTGTGTAAGTGTGTTGAACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CCGCGCTAAGGAGATTAACAGCCAGTGCATGAGTTTGGCTGTCTTCCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 DB 1141 CAGATCGCGATTAACACCGCGCTGATTAATTAACCGCGCCGCGAGAGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGCGCAGATGTGAAGAGCGGAATGCGAGGGCATGTCTGTTGGCAGGTTTACGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 DB 1261 ATCGTGAATGAAGACAGCATACCGCGCTGCATTTGAGATCTCATCACCATACCATCAC 1320
 RESULT 3
 ARI30956
 LOCUS ARI30956 1323 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 1 from patent US 6190897.
 ACCESSION ARI30956
 VERSION ARI30956.1 GI:14119281
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1323)
 AUTHORS Kretz, K.
 TITLE Phylase
 JOURNAL Patent: US 6190897-A 1 20-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..1323
 /organism="unknown"
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3, 27e-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566A-2 (1-440) x ARI30956 (1-1323)
 QY 1 MettysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGAGCATTTATCCATTTTATCTCTTCGATTCCTTAACCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGCTGTGATGTCAGTGTGATGCT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 121 GTGCTGTCTCCAAACCAAGGCCAGCACTGATGAGGATGTCAACCCAGACCAATGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLysLeuIleAlaTrpLeu 80
 DB 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCCGNGTGGTGGAGCTAATGCCATATCTC 240
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 241 GGACATTTACCAACGCCAGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGGCTGCCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTGTGTGACAGTGGCATTATTCCTGATGTGACAGACCGTACCCGTAACACAGCGCA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCCCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGCCAGATACG 420

QY 141 SerSerProaspProleuPheasnProleuLysThrGlyValCysGlnLeuaspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAACTCTAAATAACGCGGTTGGCACTGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleleuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGAGCTGACGGGATCTCAGCAGGGCAGAGGGCTAAATTCGACTTAAACGGGCAAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAAGCGGGGTTTCGGGAACCTGSAACGGGGTCTTAATTTCCGCATCAACTGTGGC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACCTGAGAAACAGACGAGAAAGCTGTTCAATACGAGGCAATTACATCGGAACTC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACCCCGACAAATGCTCATTAACCGGTGGTGAAGCTCGCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyLysArgIleThr 260
 DB 721 GAGATATTCTCTCCACAAACAGCAGGAAATGCCGAGCGGGGTGGGGAAGATCAC 780
 QY 261 AspSerHisGlnThrPasnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 781 GATTCACACAGGTGACACCTGCTAAGTTGCAATACGCGCAATTTATTTGCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCGAGAGTTGCCCGCAGCGCCGCCCACTTATTCGATTCATCGACGCG 900
 QY 301 LeuThrProHisProGluGlnLysGlnAlaThrGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATCCACCCCAAAACAGCGGATAGTGTGACATTAACCACTTCACTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnThr 340
 DB 961 TTTATTCGCGGACACGATACTAATCTGGCAAAATCTGGCGGCGCTGAGCTCAACGCG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgThr 360
 DB 1021 ACGCTCCCGGTCAACCGCGGATTAACGCGCGCGAGTGTGAACTGTGTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTAACGATTAACAGCCAGTCAAGTTCGCTGCTTCACACTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
 DB 1141 CAGATGCGTGAATAAACCGCCCTGTCATTAATACGCGCGGAGAGGTGAACCTGAC 1200
 QY 401 LeuAlaGlyCysGluGlnLysArgAsnAlaGlnLysMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGATGTGAAGACGAAATCCGACAGGCGATGTTCTGTTGCGAGGTTTTCGGA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 DB 1261 ATCGTGAATGAAGACGATACCGCGCTGCAATTCGATTCATCACCACATCACCATCAC 1320
 RESULT 4
 AX052713 LOCUS AX052713 1323 bp DNA linear PAT 12-JAN-2001
 DEFINITION Sequence 1 from Patent WO0071728.
 ACCESSION AX052713
 VERSION AX052713.1 GI:12226903
 KEYWORDS
 ORGANISM Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1 to 1323)

AUTHORS Short, J.M. and Kretz, K.A.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0071728-A 1 30-NOV-2000;
 Diversa Corporation (US)
 FEATURES location/Qualifiers
 source 1..1323
 CDS
 1..1323
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC21810.1"
 /db_xref="GI:12226904"
 /translation="MKAILIPFLSLILPLPQSAFQSEPELKLEVVYVSRHGVAP
 TKATQIMQDTPDAMPVPRVKGLMTPRGGLIAYLGHYQRLVADGLAKGCPQS
 GOVALIADVDERTRKTEGAFAGLAPDCAIVHNOADSSPDPLNPLKGVCLDPA
 NVTDAIISRAGSITADPFGHROTAPRELERLNPQSLCKRREKQDSGLTQALPS
 ELKTSADNVSITGAVSLASMLTEIFLIDQAOQMEPQRCRTTDSHONTLSLINAQF
 YLQRTPEVARSRAIPRLDLTMAALTTPHPQKQVGYVLPVSLVLEIGHDTNLNLGG
 ALELWTLPGQPDNTPPGELVFERWRLSDNSQMIQVSLVFTLOQMRDRTPLSLNT
 PPGEVKTLTLACSEERNAQMSLAGFTQIVNEARIPACSLSRSHHHHH"
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,27e-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX052713 (1-1323)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCGATTCGTTAACCCCGCAATCTCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACCGCGAGCTGAAAGTGGAGATGTCAGTCGTCATGCT 120
 QY 41 ValAlaGlnProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGCTGCTCCAAACCAAGCCAGCACTGATGCAAGATGTCACCCCGACAGCAAGGCCA 180
 QY 61 ThrTrpProValLysLeuGlyLysLeuThrProArgGlyGlyGluLeuIleAlaThrLeu 80
 DB 181 ACCTGCGCGGTRAAACTGGGTGGCTGACACCGCGNGTGTGAGCTTAATGCCATATCTC 240
 QY 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
 DB 241 GGACATTAACCAAGCCAGCTGCTGATACCGCAGGATGCTGGCGAAAGGCTGCCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGCTCAGCGTCGCTAATTCGATGCGCAGCGCTACCCCGTAAACAGCGCAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCCCGCGGCTGGCACTGACTGTGCATTAACCGTACATCCACAGCGAGATACG 420
 QY 141 SerSerProaspProleuPheasnProleuLysThrGlyValCysGlnLeuaspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAACTCTAAATAACGCGGTTGGCACTGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleleuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGAGCTGACGGGATCTCAGCAGGGCAGAGGGCTAAATTCGACTTAAACGGGCAAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnPheProGlnSerAsnLeuGly 200

Db 541 CGGCAAAACGGCGTTTCGGCAACGTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
 QY LeuysaragluylusglnaspglusercysserleuthrglnalaLeuProsergluLeu 220
 Db 601 CTTAAAGCTGAGAAACAGGAGGAAAGCTGTTCAATTACCGGCACTTACCATCGGAATC 660
 QY LysValSerAlaaspAsnValserLeuthrglyAlaValSerLeuAlaSerMetLeuthr 240
 Db 661 AAGGTAGCGCCGCAACATGCTCATTTACCGGTGGGTAAGCTTCGCATCATGTGACG 720
 QY GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProgluProglyTyrPglYArgIleThr 260
 Db 721 GAGATATTTCCTCCGCAACAGCACAGGAAATGCCGGAGCCGGGGTGGGAAGGATCAAC 780
 QY AspSerHisglnTTPAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 781 GATTCACACCACTGGACACCTTGTGTAAGTTTGCAATACGGCAATTTATTTGCTACAA 840
 QY ArgThrProgluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 841 CGCAGCCGAGAGTGGCCGCGGCGGCGGCGGCAACCCGCTTATGGATTGATCATGCGACG 900
 QY LeuthrProHisProProgluInlysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 Db 901 TTGACGCCCCATCCACCCCAAAACAGCGCATGTGACATTCACCATTCACACTACTG 960
 QY PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnThr 340
 Db 961 TTTATGCGGACACGATGACTATCTGCAAAATCTCGGCGGCGCACTGAGCTCAACTGG 1020
 QY ThrLeuProgluInProAspAsnThrProProgluGlyGluLeuValPheGluArgThr 360
 Db 1021 ACGCTCCCGGTCACGCGGATTAACACGCGCGGAGTGTGAACCTGTTTAAAGCTGG 1080
 QY ArgArgLeuSerAspAsnSerGlnTyrPheGlnValSerLeuValPheGlnThrLeuGln 380
 Db 1081 CGTCGGCTTAACGATTAACAGCCAGGATTCAGGTTTCGCTGCTCCACACTTTACAG 1140
 QY GlnMetArgAspLysThrProLeuSerLeuAsnThrProProgluGlyValLysLeuThr 400
 Db 1141 CAGATGCGGTAAACCCCGCTGCTAATTAATACGCGCGCGGAGGTGAACCTGACC 1200
 QY LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 1201 CTGCGAGGATGGAAGACGGAATGCGCAGGCACTGCTTCCTGGCAGCTTTACGCA 1260
 QY IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 Db 1261 ATCGGATGAAGCAGCATACCGGCGTGAAGTTGAGATCTCATCACCATCACCATCAC 1320
 RESULT 5
 AX338966 LOCUS AX338966 1323 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 1 from Patent WO0189317.
 ACCESSION AX338966
 VERSION AX338966.1 GI:18129102
 KEYWORDS
 ORGANISM Escherichia coli.
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1
 AUTHORS Short, J.M., Kreitz, K.A. and O'Donoghue, E.
 TITLE Dietary aids and methods of use thereof.
 JOURNAL Patent: WO 0189317-A 1 29-NOV-2001;
 DIVERSA CORPORATION (US)
 FEATURES
 source
 1. 1323
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1. 1323
 /note="unnamed protein product"
 /codon_start=1

/transl_table=1
 /protein_id="CAD20553.1"
 /db_xref="GI:18129103"
 /translation="MKATLPELILPLPQSAFAQSEPELKIESVYIRHGVRAP
 TKATQMDVTPDAMPPTTPYVKLGMITPPGSELVAYLGHQORLYABDILAKKCPQS
 GOVALIADVDERTKRTGEAFAGLAPDCAIPLVDITSSDPLEPLKTVCCODNA
 NWDAIILSRAGSILADFHROTAFRELEARNPQSMCLKRRKODESCSLTALPS
 ELKYSADVSLTGAVSLASMLTEIFLQOAGMPEPMGRITDSHOWTSLINAPF
 YLORTPEVRSRATPLILDLMAALTPEPPOKQAYGYTLPTSYLFIGHDTNLTNG
 ALLENTPLPGQPDPTPGSELVEFRERLSDNSQIMQVSYFOTLQORDKTPISLNT
 PPEVKTLIACCEERNAGKCSLQGTQIVNEARIPKCSLSHHHHH"
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 276-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX338966 (1-1323)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1 ATGAAGGAGGATCTTAATCCATTTTATCTCTGATTCGTTAACCCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 Db 61 TTCGCTCAGAGTACCCGAGCTGAAGCTGGAAGCTGCGATGTGTCAGTCGTCATGT 120
 QY 41 ValArgAlaProThrThrLysAlaThrGlnLeuMetGlnAsnPvalThrProAspAlaThrPro 60
 Db 121 GTGCGTCTCCACCAAGGCCACGACCTGAGTGCAGATGTCACCCAGACGATGGCCA 180
 QY 61 ThrTyrProValLysLeuGlyTyrPheLeuthrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 Db 181 ACCTGGCGGTAAACCTGGTGGCTGACACCCGCGNGGTGGTGCATATGCGCTATCTC 240
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 241 GGACATTTTACCAAGCGCACGCTGCTGACCCGAGATTCGCGCAAAAGCGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyLeu 120
 Db 301 CAGTGTGTCAGTGCAGATTAATCTGATATGTCAGACGAGCCGTACCCGTAAACAGCGCA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 361 GCCTTCGCGCGCGGCTGGCACCTGCTGTCATTAACCGTATACCCAGGCAAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 421 TCCAGTCCCGATCCCTTAATTAATCTCTAAAAACGTGCGTTGCCAATGATTAACCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
 Db 481 AACGTGATGACGCGCATCTTCAGCAGGAGGAGGAGGTCATATGCTTACCGGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuGly 200
 Db 541 CGGCAAAACGGCGTTTCGCGCAATGGAACGGGTCTTAATTTCCGCAATCAAACTGTGCG 600
 QY 201 LeuLysaragluylusglnaspglusercysserleuthrglnalaLeuProsergluLeu 220
 Db 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGGCACTTACCATCGGAATC 660
 QY 221 LysValSerAlaAspAsnValserLeuthrglyAlaValSerLeuAlaSerMetLeuthr 240
 Db 661 AAGGTAGCGCCGCAACATGCTCATTAACCGGTGGGTAAGCTTCGCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProgluProglyTyrPglYArgIleThr 260

```

|||||
Db 721 GAGATATTTCTCCCAACAGCAGGGAATCCGCGCGGGTGGGGAAGATCAC 780
Oy 261 ASPSERHISGLINTPASPnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGATGACACCTTGTGAAGTTTGCAATACCGCAATTTATTTGCTACAA 840
Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAla 300
Db 841 CGCAGCGCAGAGGTGGCCGCGCGCCACCCCGTTATTGATTGATCATGCGACAG 900
Oy 301 LeuThrProHisProProGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATCCACCCCAAAACAGCGGTATGATGATTAACCACTTCACATGACTCG 960
Oy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db 961 TTTATTGGCGGACACGATACTATCTGGCAAAATCTCGCGCGGCGACATGAGTCACTAC 1020
Oy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1021 ACGCTTCCCGGTACCCGCGATACACGCCCGCGAGGTGTAACCTGTGTTGAACGCTGG 1080
Oy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGGATACAGCCAGTGAATCAGGTTGCTGCTCCAGACTTACAG 1140
Oy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1141 CAGATGCGTGATAAACCGCGGTGTCATTAAATACGCCCGCGAGAGTGAATGACAC 1200
Oy 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGGATGTGAAGACGGAATCGCAGGCGCATGTGTCGTGGCAGGTTTTCGCA 1260
Oy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGAATGAAGACGATACCGCGGTGACGTTTGAGATCTCATACCATCACCATCAC 1320

RESULT 6
AX356566 1323 bp DNA linear PAT 06-FEB-2002
LOCUS AX356566
DEFINITION Sequence 1 from Patent WO0190333.
ACCESSION AX356566
VERSION AX356566.1 GI:18621053
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
          Escherichia coli.
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
REFERENCE
  1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
    O'Donoghue, E.
  Recombinant bacterial phytases and uses thereof
  Patent: WO 0190333-A 1 29-NOV-2001;
  DIVERSA CORPORATION (US)
FEATURES
  source
    1..1323
    /organism="Escherichia coli"
    /db_xref="taxon:562"
    1..1323
    /note="unnamed protein product"
    /codon_start=1
    /transl_table=11
    /protein_id="CAD23025.1"
    /db_xref="GI:18621053"
  CDS
    1..1323
    /translation="MKATILPFLSLILPVTQSAFAOSEPEIKLESVIVSRHGVRA
    PTKATQMDVTPDAMPVTPVKLGMVTPRGELIYLGXORQRLVADLILAKGCPQS
    GOVAIADVDERTRKTGEAFAGLADCAITVHQADTSSPDLPNPKTGCOLDNA
    NVTDAIISRAGSITADFTGHROTAFRELEVLNPNOSNLCKREKODSGSTQALPS
    ELKYSADVSLTGAVALASMLTEITFLQOAGMPGKRGITDTSQHWMTLSINAOE
    YLIQRTPEVASRATPLILDITMALITPPPKQAYGVTLPTSVETFIAGHDYNLANLG
    ALBNWTLPGDPDNTPEGELVFERKRLSDNSQWIOVSLVFLOQMRDXTPLSLNT

```

```

BASE COUNT 323 a 353 c 357 g 289 t 1 others
ORIGIN
Alignment Scores:
  Pred. No.: 3,27e-169
  Score: 2302.00
  Percent Similarity: 100.00%
  Best Local Similarity: 100.00%
  Query Match: 100.00%
  Gaps: 0
  DB: 6
US-09-777-566a-2 (1-440) x AX356566 (1-1323)
Oy 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGGATCTTAATCCATTTTATCTGTGATTCGGTTAAACCCGCAATCTGCA 60
Oy 21 PheAlaGlnSerGluProGluLeuLysLeuSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTACCGCGAGCTGAAGCTGGAAGTGGTGATTCAGTCGTCATGCT 120
Oy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTCTCCACACCAAGGCCACGCACTGATGAGATGTCACCCGACAGCATGGCCA 180
Oy 61 ThrTrpProValLysLeuGlyTyrPleuThrProArgLysGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTCGCGCGGTAAACTGGGTGGCTGACACCGCGNGGTGGAGCTAATGCGCTATCTC 240
Oy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTCACCAAGCCAGCAGCTGTGTGACCGCAGATTGCTGGCCAAAAGGCGTCCCG 300
Oy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTGTGCTGAGTGGCCATTATTCCTGATGCGACGAGCGACGACCCGTTAAACAGGCCAA 360
Oy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGCACCTGATGTCGAATTAACCGTACATACCAGCAGATACG 420
Oy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCCTTATTATATCCTTAATAAATGCGCTTGGCACTGATTAACCG 480
Oy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCTACGAGGCGAGAGGCTCAATTTGCTGACTTTACCGGCGAT 540
Oy 181 ArgGlnThrAlaPheArgGluLeuGlnLuarGlyValLeuAsnPheProGlnSerAsnLeuCys 200
Db 541 CGGCAACGCGGCTTCCGGAACCTGGAACGGGCTTAATTTCCGCATTAACCAACTTGGTC 600
Oy 201 LeuLysArgGluGlnAspGlnAspGlnSerLysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAAGAACAGAGCAGAAAGCTTTCATTACGACGAGCACTTCGCAATCGAATC 660
Oy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACCGCCGACAAATGCTCATTAACCGGTGGGTGAAGCTGCAATGCTGACG 720
Oy 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPGLYArgIleThr 260
Db 721 GAGATATTTCTCCCAACAGCAGGGAATCCGCGCGGGTGGGGAAGATCAC 780
Oy 261 ASPSERHISGLINTPASPnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGATGACACCTTGTGAAGTTTGCAATACCGCAATTTATTTGCTACAA 840
Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAla 300
Db 841 CGCAGCGCAGAGGTGGCCGCGCGCCACCCCGTTATTGATTGATCATGCGACAG 900

```

OY		301	LeuThrProHisProproGlnLysIleAlaTyrGlyValIThrLeuproThSerValIeu	320
Dd		901	TTTGAGCCCCATCCACGCCAAMAAGCGGTATGGTGACAATTACCACCTTCAGTACTG	960
OY		321	PheIleAlaGlyHisAspTrhAsnLeuAlaAsnLeuGlyGlyAlaIleuGluIleuAsnTrp	340
Dd		961	TTTATATGCCGGKACAGACTACTAATCTGGCAATCTCGCGGGCGCAGCTGAGCTGAATCG	1020
OY		341	ThrLeuProGlyGlnProAspAsnTrhProProGlyGlyGluLeuValIpheGluArgTrp	360
Dd		1021	ACGCTTCCCCTGCAGCCCGCATACACGCCGCCAGGTGGTGAACGTGTTGAACGCTGG	1080
OY		361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValIpheGlnThrLeuIen	380
Dd		1081	CGTCGGCTAAAGGATACACGACGAGGATACAGGTTGCTGGTCTTCCAGACTTTAACG	1140
OY		381	GlnMetArgAspLysTrhProLeuSerLeuAsnTrhProProGlyGlyValLysLeuTrh	400
Dd		1141	CAGATGCGCTGATPAAAAGCCCGCTGCTCATTAATAACGCCGCCCGGAGAGGTGAATGCAC	1200
OY		401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMecCysSerLeuAlaGlyPheTrhGln	420
Dd		1201	CTGGCAGCATGTGAAGACGCAAATGCCGAGGGCATGTTCGTTGGCAGGTTTACGCA	1260
OY		421	IleValaGlnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis	440
Dd		1261	ATCGGATGATGAAGCACGGCATACGGCGTGACGTTTGAATCTCATCACCATCACCATCAC	1320
RESULT 7				
LOCUS	ECOAPPAA	1901 bp	DNA	linear BCT 04-APR-2002
DEFINITION	Escherichia coli periplasmic phosphoanhydride phosphohydrolase			
ACCESSION	(appa) gene, complete cds.			
VERSION	M58708			
KEYWORDS	M58708.1 GI:145283			
SOURCE				
ORGANISM	Escherichia coli. Escherichia coli. Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (bases 1 to 1901)			
AUTHORS	Dassa,J., Marck,C. and Boquet,P.L.			
TITLE	The complete nucleotide sequence of the Escherichia coli gene appa reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase			
JOURNAL	J. Bacteriol. 172 (9), 5497-5500 (1990)			
MEDLINE	90368616			
PUBMED	2168385			
REFERENCE	2 (sites)			
AUTHORS	Ostalan,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van Etten,R.L.			
TITLE	Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase			
JOURNAL	J. Biol. Chem. 267 (32), 22830-22836 (1992)			
MEDLINE	93054596			
PUBMED	1429631			
REFERENCE	3 (sites)			
AUTHORS	Ostalan,K. and Van Etten,R.L.			
TITLE	Asp04 of Escherichia coli acid phosphatase is involved in leaving group protonation			
JOURNAL	J. Biol. Chem. 268 (28), 20778-20784 (1993)			
MEDLINE	94012612			
PUBMED	8407904			
FEATURES				
source	Location/Qualifiers 1..1901 /organism="Escherichia coli" /db_xref="taxon:562" 3..1522 /gene="appa" 3..8 /gene="appa" 16..108			
gene				
RBS				
CDS				

stem_loop	/gene="appa" /note="ORF" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAI2085.1" /db_xref="GI:145284" /translation="MWYLLNFVGIILMCSLSTLVVMDLRKKS" 118. .149
-35_signal	/gene="appa" 133. .138
-10_signal	/gene="appa" 156. .161
RBS	/gene="appa" 172. .177
CDS	/gene="appa" 188. .1486 /gene="appa" /standard_name="acid phosphatase" /codon_start=1 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AAI2086.1" /db_xref="GI:145285" /translation="KKATILIPFLSLILPITPQSAFQASEPEPKIESVIYISRRICVAP TKATQMDVPEPADMPVPRKGLMTPRGGLIAYIDHYQKQRLVAGLAKKQCPQS GGVALIADVDETRKTGEAFAGLAPDCAITVHQADTSSPDPLNPKTGVCOLDNA NTVDAILSRAGSIAIDFTGHROTAFRELERYLNFPOSMLCKREKQESCSLQAQPS ELKSAVDNVLGVAVSLASMLREIFILQQAQGMPEPGMGRTDSHOKNTLLSHNOF YILQTPYARSRATPLDLIKTKATLTPPPPKQAKGVTLPRSYVFIAGHPTNLANLGG ALELWTLLEGQPDNTPGGELVFERWRSLSDNSQWIOVSLVFTQLQOMRKDTPLSLNT PIEGEVKLTLAGCEERNAGMCSLAGFTQIVNEARIIPACSL" 188. .253
sig_peptide	/gene="appa" 254. .1483
mat_peptide	/gene="appa" /product="periplasmic phosphoanhydride phosphohydrolase" 299. .301
variation	/gene="appa" /standard_name="H303A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R16A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R20A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R92A mutant" /note="created by site directed mutagenesis" /citation=[2] 302. .304
variation	/gene="appa" /standard_name="H17N mutant" /note="created by site directed mutagenesis" /citation=[2] 1163. .1165
variation	/gene="appa" /standard_name="D304A mutant" /note="created by site directed mutagenesis" /citation=[3] /phenotype="Asp 304 changed to Ala" replace="gct" 1163. .1165
variation	/gene="appa" /standard_name="D304E mutant" /note="created by site directed mutagenesis"

/citation=[3]
 /phenotype="Asp 304 changed to Glu"
 /replacer="gag"
 1497. .1522
 /gene="appa"
 /note="rho-independent terminator"

BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 7,71e-164 Length: 1901
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 1 Gaps: 0

US-09-777-566a-2 (1-440) x ECOAPPAA (1-1901)

OY 1 MetlysalaileuileleuphelleuSerleuileleProleuthrProglInserala 20
 Db. 188 ATGAAGAGCATCTTAATCCATTTTATCTTCGTGATTCGGTTAACCCGCAATCTGCA 247
 OY 21 PhealaginSerGluProgluLeuLysleuLysSerValIleValSerArgHisGly 40
 Db 248 TTCCTCAGAGTGAGCCGAGCTGAAGCTGAAGTGTGTGATGTCAGTCGTATGT 307
 OY 41 ValArgalaProThrLysalaThrGlnleuMetGlnaSerValThrProaspAlatPro 60
 Db 308 GTGGGTGCTCCACCAAGCCAGCCAGCACTGATGCGAGTGCACCCGACGCAATCGCA 367
 OY 61 ThrTrpProValLysleuGlyTrpleuthrProArgIlyGlyLuleuIleAlaTryleu 80
 Db 368 ACCGTGCGGTAAACTGGTGTGGCTGACCGCGGTGGTGAAGTATGCGCTATCTC 427
 OY 81 GlyHsTyrgInaArgGlnaArgleuValaIaaspGlyleuAlaIyaleGlyysPro 100
 Db 428 GGACATTACCAACGCGACGCTGTGTAGCCAGGATGTGGCAAAAGGGCTGCCG 487
 OY 101 GlnSerGlyGlnValaIleIleIleIleIleIleIleIleIleIleIleIleIleIle 120
 Db 488 CAGTCTGTGTCAGTCCGATATTGCTGATGTCGACGAGCGTACCCGTAACCAAGCGCA 547
 OY 121 AlaPhealaIaGlyleuAlaProaspGlyalaIleThrValHsThrGlnAlaAspThr 140
 Db 548 GCCTTCGCGCGCGGCTGCGACCTGACTGCAATTAACCTACATACCCAGGACATACG 607
 OY 141 SerSerProaspProleuPheasnProleuLysThrGlyValGylInleuAspAla 160
 Db 608 TCCAGTCCCGATCCGTTATTTAACTCTTAATAAACTGGCTTGGCACTGGATTAACCG 667
 OY 161 AsnValThrAspAlaIleleuSerArgIaGlySerIleAlaAspPheThrGlyHis 180
 Db 668 AACGTGATGACGCGATCTCAGCAGGCGAGGAGGTCAATTCGACTTTAACCGGCGAT 727
 OY 181 ArgGlnThrAlaPheArgGluLeuGlnaArgValLeuasnPheProGlnSerasnLeuGly 200
 Db 728 CGGCAAAAGCGGTTCCGCAACTGGAAGGGTGTATTTCCGCAATCAAACTTGTGC 787
 OY 201 LeuysArgGluLysGlnaSpGlnSerGlySerleuthrGlnAlaLeuProSerGluLeu 220
 Db 788 CTTAAACGTGAGAAACAGGAGAAAGCTGTTCAATTAAAGCAGGCAATTCAGGAACTC 847
 OY 221 LysValSerAlaAspAsnValSerleuthrGlyAlaValSerleuAlaSerMetleuthr 240
 Db 848 AAGGTGACCGCGCAATGTCTCATTAACCGGTGGGTAGGCTGCTCATGCTGACG 907
 OY 241 GluIlePheleuLeuGlnaGlnaGlnaGlyMetProgluProgluTyrgIleArgIleThr 260
 Db 908 GAGATATTTCTCTGCAACAGCAGGAGATGCGGAGCGGGGTGGGAAGATCAC 967
 OY 261 AspSerHisGlnTrpAsnThrleuLeuSerleuHisAsnAlaGlnPheTyrlleuGln 280
 |||||||

Db 968 GATTCACACAGGTGACACCTTGCTTAAGTTGCATTAACGCCCAATTTTATTCCTACAA 1027
 OY 281 ArgThrProgluValaIaArgSerArgAlaThrProleuLeuAspIleuIleMetAla 300
 Db 1028 CGCAGCGCAGAGGTGGCCGCGACGCGCCACCCGTTATTAAGATTGATCAAGACAGCG 1087
 OY 301 LeuthrProHisProProgluLysGlnAlaTyrgIyValIthrleuProThrSerValleu 320
 Db 1088 TTGAGCGCCCATCCACCCCAAAACAGCGCTATGTGTGACATTAACCACTTCAGTGTG 1147
 OY 321 PheIleAlaGlyHsAspThrAsnleuAlaasnleuGlyGlyAlaLeuGluLeuasnTrp 340
 Db 1148 TTTATCCGCGGACACGATTAATCTGCAATTCGCGCGCGCACTGAGCTCACTACG 1207
 OY 341 ThrleuProgluGlnProaspAsnThrProProgluGlyGluLeuValPheGluArgTrp 360
 Db 1208 ACGTTCGCGGTGACCGCGGATTAACCGCGCGAGGTGTGAACGTGTTGAACGCTGCG 1267
 OY 361 ArgArgleuSerAspAsnSerGlnTrpIleGlnValSerleuValPheGlnThrleuGln 380
 Db 1268 CGTCGCTAACGATTAACAGCAGGAGTTCAGGTTGCTGCTTCACAGCTTTACAG 1327
 OY 381 GlnMetArgAspLysThrProleuSerleuasnThrProProgluGlyValLysleuthr 400
 Db 1328 CAGATGCTGATTAACCCCGCTGTCAATTAATACCGCGCGGAGAGTGAACTGAC 1387
 OY 401 LeuAlaGlyCyGluGluArgAsnAlaGlnaGlyMetCysSerleuAlaGlyPheThrGln 420
 Db 1388 CTGCGAGATGTGAAGACGAAATCGCAGGCGATGTGTGTTGGCAGGTTTACGCA 1447
 OY 421 IleValaGlnGluAlaArgIleProAlaLysSerleu 432
 Db 1448 ATCGTAATGAAGCAGCATACCGCGGTGCGAGTTTG 1483
 RESULT 8
 AX356572 1901 bp DNA linear PAT 06-FEB-2002
 LOCUS AX356572 Sequence 7 from Patent WO0190333.
 DEFINITION AX356572
 ACCESSION AX356572
 VERSION AX356572.1 GI:18621059
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
 O'Donoghue, E.
 Recombinant bacterial phytases and uses thereof
 Patent: WO 0190333-A 7 29-NOV-2001;
 DIVERSA CORPORATION (US)
 FEATURES
 source
 1. 1901
 Location/Qualifiers
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,71e-164 Length: 1901
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX356572 (1-1901)
 OY 1 MetlysalaileuileleuphelleuSerleuileleProleuthrProglInserala 20
 Db 188 ATGAAGAGCATCTTAATCCATTTTATCTTCGTGATTCGGTTAACCCGCAATCTGCA 247
 OY 21 PhealaginSerGluProgluLeuLysleuLysSerValIleValSerArgHisGly 40
 |||||||

```

Db      248  TTCCCTCAGAGTGGCCGACGTGAGCTGGAAGTGTGATGTGTCAGTCGTACGT 307
Qy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      308  GTGGGTGTCCCAACCAAGCCAGCACTGATGCGAGTGTCCACCCAGCAGCAGGCCA 367
Qy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80
Db      368  ACCTGGCCGTAACCAACTGGTGTGGCTGACCCGCGNGGTGGTGAAGTAAATGCTATCTC 427
Qy      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      428  GGACATTTACCAACCCAGCCGCTGTGTACCCAGCAGGATGTGGCGAAAGGGCTGCCCG 487
Qy      101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgGlyThrArgLysThrGlyL 120
Db      488  CAGTCTGTGTCAGTGGCGCATTTATTCCTTAAACTGGCGTTGGCACTGGATTAACGCG 547
Qy      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548  GCCTTCCGCCCGCCGCTGGCACCTGCTGCAATACCTACATACCCAGCAGCATACG 607
Qy      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608  TCCAGTCCCGCATCCGTTATTTAACTCTTAAACTGGCGTTGGCACTGGATTAACGCG 667
Qy      161  AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      668  AACGTGATGACGGGATCTCAGCAGGAGGAGGTCGAATGTCGACTTACCGGGCAT 727
Qy      181  ArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnProGlnSerAsnLeuGly 200
Db      728  CGCGAACAGCGGTTTCGCGAATGAGACGGGTCTTATTTTCCGCAATCAACTGTGTC 787
Qy      201  LeuLysArgGlyLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db      788  CTTAAGCTGAGAAACAGACGAAAGCTGTTCATTAACGACGATACATTCGGAATC 847
Qy      221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848  AAGGTGAGCGCCGCAATGTCTCATTAACCGGTGGGTGAACCTCTGCAATGCTGACG 907
Qy      241  GlnIlePheLeuLeuGlnGlnAlaGlnLysMetProGlnProGlyTrpGlyArgIleThr 260
Db      908  GAGATATTTCCTCGCAACACAGCAGGAGATCCCGGAGCCGGGTGGGAAGATCAAC 967
Qy      261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db      968  GATTACACACAGTGGAAACACCTTGTCTTAAGTTTGCATACGCGCAATTTTATTTGCTACA 1027
Qy      281  ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db      1028  CGCAGCGCAGAGGTGGCCGACGCGCCGCCCGCTTATTAAGATTGATTAAGACACAGC 1087
Qy      301  LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db      1088  TTGAGCGCCCATCCACCGCAAAACAGCGATATGGTGTGACATTAACCCACTTCAGTGTG 1147
Qy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db      1148  TTTATTCGCGCAGACAGTACTTAATCTGCAATCTCGCGCGCGCACTGAGACTCAACTGG 1207
Qy      341  ThrLeuProGlnLysProAspAsnThrProProGlnGlyGlyLeuValPheGlnAlaArg 360
Db      1208  ACCTTCCCGGTACCGCGGATTAACCGCGCAGGAGTGTGAATGTGTTTGAACGCGTG 1267
Qy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268  CGTGGGCTAAAGCATTAACAGCAGTGTGATTCAGTGTGCTGTGCTTCCAGACTTTAAG 1327
Qy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlnGlyValLysLeuThr 400

```

```

Db      1388  CAGATCGGTGATTAACCAACCGCTGTCAATTAATACCCGCCGAGAGGTGAACCTGACC 1387
Qy      401  LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388  CTGGCAGAGTGTGAAGACAGGAATGCGCAGGCACTGTGTGCTTGGCAGGTTTACGCA 1447
Qy      421  IleValAsnGlnValAlaArgIleProAlaCysSerLeu 432
Db      1448  ATCGTGAATGAAGCAGCATACCGCGCTGACGTTTG 1483

RESULT 9
AX042376          3470 bp   DNA       linear   PAT 23-NOV-2000
LOCUS             Sequence 5 from Patent WO0064247.
DEFINITION        AX042376
ACCESSION         AX042376.1 GI:11340994
VERSION           AX042376.1
KEYWORDS           synthetic construct.
SOURCE             synthetic construct.
ORGANISM           artificial sequences.
REFERENCE          1 (bases 1 to 3470)
AUTHORS            Forsberg,C.W., Golovan,S. and Phillips,J.P.
TITLE              Transgenic animals expressing salivary proteins
JOURNAL            Patent: WO 0064247-A 5 02-NOV-2000;
                    UNIVERSITY OF GUELPH (CA)
FEATURES           Location/Qualifiers
                     source          1..3470
                                     /organism="synthetic construct"
                                     /db_xref="taxon:32630"
                                     /note="R15/APPA transgene"
BASE COUNT         1065 a      721 c      735 g      949 t
ORIGIN
Alignment Scores:
Pred. No.:         1,57e-163      Length:      3470
Score:             2235.00         Matches:     430
Percent Similarity: 99.54%         Conservative: 0
Best Local Similarity: 99.54%      Mismatches:  2
Query Match:       97.09%          Indels:       0
DB:                6              Gaps:          0

US-09-777-566a-2 (1-440) x AX042376 (1-3470)
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db      1811  ATGAAGGATCTTAATCCATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 1870
Qy      21  PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db      1871  TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGTGTGATGTGTCAGTCGTACGTGT 1930
Qy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      1931  GTGCTGTCTCCAAACCAAGCCAGCACTGATGACGAGTGTCCACCCAGACGATGGCCA 1990
Qy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80
Db      1991  ACCTGGCCGTAACCAACTGGTGTGGCTGACACCGCGGTGTGAGCTAATGCGCTATCTC 2050
Qy      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      2051  GGACATTTACCAACGCGCAGCTGTGTACCCGACGATTTGGCGCAAAAGGGCTGCCCG 2110
Qy      101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgGlyThrArgLysThrGlyL 120
Db      2111  CAGTCTGTGTCAGTGGCGCATTTATTCCTATGTCGACGAGCGGTACCCGTAACACAGCGAA 2170
Qy      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      2171  GCCTTCGCCCGCGGCTGGCACCTGCTGCAATTAACCTACATACCCAGCAGATACG 2230
Qy      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

```

Db 2231 TCACATCCCGATCCGTTATTATATCTCTAAATAACGCGCTTGGCAACGTGATTAACGG 2290
 Qy 161 AsnValThrAspAlaIleuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 2291 AACGACTGACGCGATCTCGACAGGAGGAGGCTCAATGCTGACTTTACCGGGCAT 2350
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuGly 200
 Db 2351 CGGCAACGGCGCTTCCGGAACGTGACGGGCTGTTAATTTTCCCAATCAAACTTGTC 2410
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerIleuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAAACGTGAAACAGACGAAAGCTGTTCAATACGACGACCTTACCTCGAATC 2470
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTAGACCGCCGCAATATGTCATTAAACGGGTGCGTAAAGCTGCATCAATGCTGACG 2530
 Qy 241 GluIlePheLeuLeuGlnAlaGlnAlaGlnMetProGluProGluProGluArgIleThr 260
 Db 2531 GAGATATTTCTCTCCCAACAGCACAGGGAATGCCGAGCCGGGGTGGGAAGATCAC 2590
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 2591 GATTCAACACAGTGAACACCTTGCTAAGTTTGCTAAACGCCCAATTTATTGCTACAA 2650
 Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
 Db 2651 CGCAGCGCAGAGGTGGCCGCGACCGCGCCGCTTATGATTTGATTCAGAACAGCG 2710
 Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 Db 2711 TTGACGCCCCATCCACCCCAAAACAGCGGATGAGTGACATTACCACTTCACATGCTG 2770
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaIleuGluLeuAsnThr 340
 Db 2771 TTTATCCGCGGACACAGATACATCTGCGCAATCTCGCGCGGCGACCTGAGCTCACTG 2830
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
 Db 2831 ACGCTTCCCGGTCAACCCGATACACGCCCGCGAGTGATGAACGTGTGTAACGCTG 2890
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 2891 CGTGGCGTAAACGATACAGCAGGATGAGTTCAGGTTGCTGCTTCAGACTTTAAG 2950
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValIleLysLeuThr 400
 Db 2951 CAGATGCGTGAATAAACGCCCTGTCATTAAATACGCCGCCGCGAGAGGTGAACCTGAC 3010
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGCAGAGATGTGAAGACGAAATCGCAGGAGCATGTGCTGTCGACAGGTTTACGCA 3070
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGTAATGAAGACACGATACCGCTTCAGGTTTG 3106
 RESULT 10
 AX042374 4060 bp DNA linear PAT 23-NOV-2000
 LOCUS AX042374
 DEFINITION Sequence 3 from Patent WO0064247.
 ACCESSION AX042374
 VERSION AX042374.1 GI:11340992
 KEYWORDS
 ORGANISM
 SOURCE
 synthetic construct.
 artificial sequences.
 REFERENCES
 1 (bases 1 to 4060)
 Forsberg, C.W., Golovan, S. and Phillips, J.P.
 Transgenic animals expressing salivary proteins
 Patent: WO 0064247-A 3 02-NOV-2000;
 UNIVERSITY OF GUELPH (CA)

FEATURES
 source 1..4060 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="R15/APPA + Intron transgene"
 BASE COUNT 1257 a 814 c 843 g 1146 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,89e-163 Length: 4060
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: Gaps: 0
 US-09-777-566a-2 (1-440) x AX042374 (1-4060)
 Qy 1 MetLysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1811 ATGAAGCCATCTTAATCCATTATCTCTGATTCCTTAACCCGCAATCTGCA 1870
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 Db 1871 TTGCTCAGAGTGAAGCCGAGCTGAAGTGAAGTGTGATGTTGCTGCTGCTGCT 1930
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 1931 GTGGCTGCTCCAAACAGCCAGCCAGCACTGATGAGAGATGTCAACCCAGACGATGGCA 1990
 Qy 61 ThrTrpProValLysLeuGlyTyrPheLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 Db 1991 ACCCTGCGCGTAAACTGGGTGGCTGACACCGCGGGGTGGACTAATCCCTATCTC 2050
 Qy 81 GlyHisTyrGlnArgGluArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 2051 GGACATTCACACGCCAGCGCTTGATGAGCCGACGATTCCTGCGAAAAAGGCGTCCCG 2110
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 2111 CAGCTGGTCAAGTGGCGATTTATGCTGATGTCGACGAGCGATACCCGTTAAACAGCGCA 2170
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 2171 GCCTTGCGCGCGGGCTGGCACCCTGATGTAATACCGTCAATACCCAGGACAGATAGC 2230
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 2231 TCACATCCCGATCCGTTATTATCTCTAAATAACGCGCTTGGCAACGTGATTAACGG 2290
 Qy 161 AsnValThrAspAlaIleuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 2291 AACGACTGACGCGATCTCGACAGGAGGAGGCTCAATGCTGACTTTACCGGGCAT 2350
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuGly 200
 Db 2351 CGGCAACGGCGCTTCCGGAACGTGACGGGCTGTTAATTTTCCCAATCAAACTTGTC 2410
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerIleuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAAACGTGAAACAGACGAAAGCTGTTCAATACGACGACCTTACCTCGAATC 2470
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTAGACCGCCGCAATATGTCATTAAACGGGTGCGTAAAGCTGCATCAATGCTGACG 2530
 Qy 241 GluIlePheLeuLeuGlnAlaGlnAlaGlnMetProGluProGluProGluArgIleThr 260
 Db 2531 GAGATATTTCTCTCCCAACAGCACAGGGAATGCCGAGCCGGGGTGGGAAGATCAC 2590
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 2591 GATTCAACACAGTGAACACCTTGCTAAGTTTGCTAAACGCCCAATTTATTGCTACAA 2650

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
|||||
Db 2651 CGCAGCCGACAGGTTGCCCCGACGCCGCCACCCCGTTATGATTTATATCAAGACAGCG 2710
QY 301 LeuThrProHisProProGluInLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
2711 TTGAGGCCGCCATCCACGCCAAACAGCGGTATGGTGTGACATTACCACTTCAAGTGGCTG 2770
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
|||||
Db 2771 TTTATCCGCCGACACGATACCTAATCTGTGCAATCTCGCGCGGACACTGAGCTCAACTGG 2830
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAlaThrTrp 360
2831 ACGCTTCCCGGTCCAGCGGATTAACCGCCGCGAGGTGTGAACGTGTGTTGAACGCTGG 2890
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
2891 CGTCGGCTAAGCATTAACAGCCAGATCGATTCAGGTTTGGTCTTCCAGACTTTACAG 2950
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 2951 CAGATGCCGTGATAAACGCCGCTGTCAATTAATACGCCGCCGAGAGGTGAACCTGACC 3010
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 3011 CTGGCAGGATGTGAAGACCGAAATGCCAGGGCATGTGTTCGTTGGCAGGTTTTCACCA 3070
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 3071 ATCGTGATGAAGCAGCATACCGCTTCGACGTTTG 3106

RESULT 11
AX042375 6116 bp DNA linear PAT 23-NOV-2000
LOCUS AX042375
DEFINITION Sequence 4 from Patent WO0064247.
ACCESSION AX042375
VERSION AX042375.1 GI:11340993
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 6116)
AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 4 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
source location/Qualifiers
1..6116
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="R15/APRA plasmid with pBLCAT3 vector"

BASE COUNT 1724 a 1386 c 1407 g 1599 t
ORIGIN
Alignment Scores:
Pred. No.: 3 06e-163 Length: 6116
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
Gaps: 0
US-09-777-566a-2 (1-440) x AX042375 (1-6116)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1811 ATGAAGGCATCTTAATCCATTATCTCTTCGATTCGCTTAACCCGCAATCTGCA 1870
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleIleValSerArgHisGly 40
Db 1871 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGCTGTGTGATGTGTCAGTGTGATGCT 1930

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
|||||
Db 1931 GTGGCTGCTCCACCAACGACGACGACGATGTCACCCGACGACGACGACGACGACGAC 1990
QY 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
1991 ACGTGGCGGTAACCTGGGTGGGTGACACCGCGCGGTGGTGAAGCTAATGCGCTATCTC 2050
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
2051 GGACATTTACCAACGCGACCGTGTGTACCCACAGGATTTCTGGCGAAAGGCGCGCCG 2110
QY 101 GlnSerGlyGlnValAlaIleIleLeuAspValAspGluArgThrArgLysThrGlyGlu 120
2111 CAGTGTGTGACAGGTGCGCATTTATGTGATGTGACAGCGGATACCCGTAACAGCGCA 2170
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
2171 GCCTTCCGCGCGGCGGCTGGACACTGTGCAATTAACCGTACATACCCAGCAGATACG 2230
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
2231 TCCAGTCCCGATCCGTTATTAATCTTAATAACCTGGCGTTGCCAACTGATTAACGCG 2290
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
2231 AACGTCAGTACGCGCATCTTCAGCGAGGCGAGGAGGTCAATGTGCTATTAACCGGAT 2350
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuGly 200
2351 CGGCAACGCGCGTTTCGGAACCTGGAACGCGGCTTAATTTCCCAATCAAACTGTGTC 2410
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
2411 CTTAAACGTGAGAAACAGACGAAACGCTTCAATTAACGACGATTAACCACTGGAATGC 2470
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
2471 AAGGTGACGCGCGACAAATGTCATTAACCGGTGGGTGAAGCTGCAATGCTGACG 2530
QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPheLysArgIleThr 260
2531 GAGATATATTTCTCTCAACAGACAGGAAATCCGAGACCGGGGTGGGAAGATACACC 2590
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
2591 GATTTCACACACAGTGAACACCTGTGTAAGTTGCAATACGCGCAATTTATTTGCTACAA 2650
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
2651 CGCAGCCGACAGGTTGCCCCGACGCCGCCACCCCGTTATGATTTATATCAAGACAGCG 2710
QY 301 LeuThrProHisProProGluInLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
2711 TTGAGGCCGCCATCCACGCCAAACAGCGGTATGGTGTGACATTACCACTTCAAGTGGCTG 2770
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
2771 TTTATCCGCCGACACGATACCTAATCTGTGCAATCTCGCGCGGACACTGAGCTCAACTGG 2830
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAlaThrTrp 360
2831 ACGCTTCCCGGTCCAGCGGATTAACCGCCGCGAGGTGTGAACGTGTGTTGAACGCTGG 2890
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
2891 CGTCGGCTAAGCATTAACAGCCAGATTCAGGTTTGGTCTTCCAGACTTTACAG 2950
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
2951 CAGATGCCGTGATAAACGCCGCTGTCAATTAATTAACGCCGCCGAGAGGTGAACCTGACC 3010

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGAGAGATGTGAAGACGGAATATCGAGGCGCATGTCTCGTTCAGAGTTTACGAA 3070
 QY 421 IleValaGlnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGATGAATGAAGCAGCATACCCGCTTGACGATTGG 3106

RESULT 12
 AX042373 6708 bp DNA Linear PAT 23-NOV-2000
 LOCUS Sequence 2 from Patent WO0064247.
 DEFINITION AX042373
 ACCESSION AX042373
 VERSION AX042373.1 GI:11340991
 KEYWORDS
 ORGANISM
 SOURCE
 synthetic construct.
 artificial sequences.
 1 (bases 1 to 6708)
 REFERENCE
 AUTHORS Forsberg,C.W., Golovan,S. and Phillips,J.P.
 TITLE Transgenic animals expressing salivary proteins
 JOURNAL Patent: WO 0064247-A 2 02-NOV-2000;
 UNIVERSITY OF GUELPH (CA)
 FEATURES
 source
 1. 6708
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="R15/APPA + Intron plasmid with pBICAN3 vector"
 BASE COUNT 1916 a 1479 c 1515 g 1798 t
 ORIGIN

Alignment Scores:
 Pred. No.: 3.42e-163 Length: 6708
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 Gaps: 0

US-09-777-566a-2 (1-440) x AX042373 (1-6708)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 Db 1811 ATGAAGCATCTTAATCCATTTTATCTCTTCGATTCGCTTAACCCCGCAATCTCA 1870
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 Db 1871 TTCCTCTCAGAGTGAAGCCGAGCTGAAAGCTGGAAGTGTGATGTCAGTCGTCATGGT 1930
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 Db 1931 GTGGCTGCTCCACACCAAGCCACGCACTGATGCGAGATGTCACCCACAGCGCATGGCCA 1990
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 Db 1991 ACCTGGCCGGTAAACTGGTGGCTGCACACCGCGGTGGTGAAGCTAATGCGCATCTC 2050
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 2051 GGACATTATACCAACCCACCGCTCTGTGACCGACGATGTCGGGAAAAAGGCGTGGCCG 2110
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
 Db 2111 CAGTCTGTGTCAGGTCGCGCATTTATGCTGATGTCGACGAGCTACCCGTTAAACAGCGCAA 2170
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 2171 GCGTTCGGCGCGGGGTGGCGACCTGACGTGTCATTAACCGTACATACCCAGCATACG 2230
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 2231 TTCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTTTGCACAACTGGATTAACCGC 2290

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlnHis 180
 Db 2291 AACGTGACTGACGCGCATCTTCAGCAGAGAGAGGATCAATTGCTGACTTTACCGGCAT 2350
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnIuArgValIleuAsnPhleProGlnSerLysLeu 200
 Db 2351 CGGCAAGCGGCTTTCGCAACTGGACGGGTGCTTAATTTTCCCAATCAAACTGTTC 2410
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCATTAAACGACGATTTACCATCGGAAC 2470
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGTGAAGCTTCATCAATGCTGACG 2530
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyLysArgIleThr 260
 Db 2531 GAGATATTTCTCTCCAAACAGACAGGGAATGCGGAGCGGGGTGGGAAGATCAC 2590
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 Db 2591 GATTCACACAGGTGGAACACCTTGCTAGTTGATTAACGCGCATTTTATTTGCTACAA 2650
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 2651 CGCAGCCGAGAGGTTCGCCGCGCCGACCCCGCTTATGATTTGATTCATCAAGACACG 2710
 QY 301 LeuThrProHisProProGluGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 Db 2711 TTGACGCCCATCCACCCGCAAAACAGCGTATGCTGATTAACCATTCATCACTGCTG 2770
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 Db 2771 TTTATCGCCGCGACAGATACATCATCTGCAAAATCTCGCGCGGCACTGAGTCAACTCG 2830
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 2831 ACGCTTCCCGGACCCGATCAACACGCGCGAGGTGTGAACTGTGTTGAAACGCTGG 2890
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 2891 CGTGCGCTAAGCGATTAACGACGATGATTCAGTTCAGTTTGGCTTTCACAGACTTAAG 2950
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 Db 2951 CAGATCGCTGATTAACCCCGCTGCTCATTAATACGCGCGCGAGAGTGAACACTGACC 3010
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGCAGATGTGAAGACGGAATGCGAGGCGCATGTCTGCTGGCAGGTTTAACGCA 3070
 QY 421 IleValaGlnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGATGAATGAAGCAGCATACCCGCTTGACGATTGG 3106

RESULT 13
 AE000200 10784 bp DNA Linear BCT 01-DEC-2000
 LOCUS Escherichia coli K12 MG1655 section 90 of 400 of the complete genome.
 DEFINITION AE000200 U00096
 ACCESSION AE000200.1 GI:2367111
 VERSION AE000200.1 GI:2367111
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli K12.
 Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae;
 Escherichia.
 1 (bases 1 to 10784)
 REFERENCE
 AUTHORS Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
 Riley,M., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
 Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
 Mau,B. and Shao,Y.

TITLE The complete genome sequence of *Escherichia coli* K-12
 JOURNAL Science 277 (5311), 1453-1474 (1997)
 MEDLINE 97426617
 PUBMED 9278503
 REFERENCE 2 (bases 1 to 10784)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 3 (bases 1 to 10784)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 4 (bases 1 to 10784)
 AUTHORS Plunkett, G. III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 On Sep 9, 1997 this sequence version replaced gi:1787214.
 This sequence was determined by the E. coli Genome Project at the
 University of Wisconsin-Madison (Frederick R. Blattner, director).
 Supported by NIH grants HG00301 and HG01428 (from the Human Genome
 Project and NCHGR). The entire sequence was independently
 determined from E. coli K12 strain MG1655. Predicted open reading
 frames were determined using Genemark software, kindly supplied by
 Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
 have been correlated with genetic loci are being annotated with CG
 Site Nos., unique ID nos. for the genes in the E. coli Genetic
 Stock Center (CGSC) database at Yale University, kindly supplied by
 Mary Berlyn. A public version of the database is accessible
 (http://cgsc.biology.yale.edu). Annotation of the genome is an
 ongoing task whose goal is to make the genome sequence more useful
 by correlating it with other data. Comments to the authors are
 appreciated. Updated information will be available at the E. coli
 Genome Project's World Wide Web site
 (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
 its annotations are periodically updated; this is version M54. No
 sequence changes. Annotation updates: updated gene identifications
 and products; all new functional assignments courtesy of Monica
 Riley; added promoters, protein binding sites, and repeated
 sequences described in reference 1. The unique numeric identifiers
 beginning with a lowercase 'b' assigned to each gene (protein- or
 RNA-encoding) are now designated as gene synonyms instead of
 labels. This should allow them to be searched for in Entrez as gene
 names.

FEATURES
 source 1. Location/Qualifiers
 1. 10784 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="MG1655"
 /db_xref="taxon:83333"
 49. 84 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 123. 1421 /gene="appa"
 /note="b0980"
 123. 1421 /gene="appa"
 /EC_number="3.1.3.2"
 /EC_number="3.1.3.26"
 /function="enzyme: Central intermediary metabolism. Pool,
 multipurpose conversions"
 /note="0432; 100 pct identical to PPA_ECOLI SW: P07102"
 /codon_start=1
 /transl_table=1
 /product="phosphoanhydride phosphorylase; pH 2.5 acid

phosphatase: periplasmic"
 /protein_id="AAC74065.1"
 /db_xref="GI:1787215"
 /translation="MKAILIFPLSLILPLTPQSAFPAQSEPELKIESVYISRCGVAP
 TKATQIMQDVTPMPTPVKILGWLTPRGGLIAYLGHYQQRILVADGLLAKKCCPS
 GOVALIADVDERTKGTGEAFAGLAPCAITVHQADTSSPDPFNFLKTKGCOLDA
 NVTDAIISRAGSIADEFGHROTFRLEERLVNPOSNCLIKRKOESCSLTQALS
 ELKASADVNSLSTGAVSLASMTETFLIQAGQMPGPGMGRTSDSHMRNTLISNAQF
 YLIDRTPEVASRSATPLDLIKTLTLTHPPPKQAKYGTPLPSVYFIGHDTNLNLG
 ALLENMTLPGQDTPPGGLVPEFRMRSLSDNSQWISLVFQTLQMDRDTPLSLNT
 PGPEVKTLTLAGCEERNMAQMGCSLAFQIYVEARIIPACSL"
 1153. 1189 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 1491. 1524 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 complement(11536. 3716)
 /gene="yccc"
 /note="b0981"
 complement(1536. 3716)
 /gene="yccc"
 /function="orf: Unknown"
 /note="f726; 100 pct identical to fragment YCCC_ECOLI
 SW:P38134 but has 627 additional N-terminal aa and 2 less
 C-terminal residues"
 /codon_start=1
 /transl_table=1
 /product="orf, hypothetical protein"
 /protein_id="AAC74066.1"
 /db_xref="GI:1787216"
 /translation="MTTKNMTTPGSGTQENEDILRLVGLMDHRRKFLISTATFTLI
 AVASLSTPIYQADTVQVQKGNMILSTSMINSSESPELIQOLSLTICK
 TIALNLNDIYEQKTFPIVGRKARILKREKELAISNMHLPQNGDQQLTVGEN
 GHYTLDEEFTVQNVQGVQERKEDVALTADIKAKPKQFVLSQRTLEALNALQETP
 TVSSRSKESGMLKLTWGDQDLITRLINSANNYLQONLNRAQDOSQSEFLQRL
 PVESELDQAEKINLVYRQODSDVNLNEAKAVLEQIYVYNQNLNLTFRREISQY
 KKHPTYRALKEGRTLEQERKRLNKRVSAMPSTQOEYLRSLRQVEGAVYLLQMLR
 QOETISKSSAIGNVRITIDPAVPOPOVKKKALNVYGLFGLFISGAVTALAMR
 RGVAPRQLEHGISVYATTPMSMDKRTLRKKNLFSNQQRRTKNIPPLANDNA
 DSAVEAVRALTSLHFMETENNIMLTGATPDSGTFVSTTAAVYAQSDQVLT
 DADIRGSHNLFTVSNHEGLSEYLAGDEKNTVIOHFGKGFVIRGVQPPNSEL
 LMRMRQLELMADHYDLVYDPPMLAVSDAIVAGSVGTSLSLVARFGLNTAKES
 LSMQRDQAGVIGKIALINGVYIKRASAVSYGVYVYSYSEKE"
 complement(3733. 3764)
 /note="factor Sigma70; predicted +1 start at 1043443"
 complement(3736. 4194)
 /gene="yccy"
 /note="b0982"
 complement(3736. 4194)
 /gene="yccy"
 /function="putative enzyme; Not classified"
 /note="f152; This 152 aa ORF is 33 pct identical (13 gaps)
 to 139 residues of an approx. 160 aa protein PPA_L_SCHPO
 SW: P41893"
 /codon_start=1
 /transl_table=1
 /product="putative phosphatase"
 /protein_id="AAC74067.1"
 /db_xref="GI:1787217"
 /translation="MDSNMAQLKFNLSILVYCTGICRSPIGERLLRRLPGVKYSAG
 VHGLVKHPADATADVANAHGVSLSEGHGRGLTPEMARNDOLITLAMESEHIAQGTATA
 PEVGRKTMLEQGMLEQKEIPPRYKRSQDAFHYVGMLEKASQENAKRLSR"
 complement(3793. 3820)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043503"
 complement(3818. 3844)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043528"
 complement(3865. 3892)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043575"
 complement(4170. 5309)
 gene

CDS
 /gene="ycsc"
 /note="b0983"
 complement(4170..5309)
 /gene="ycsc"
 /function="phenotype: Not classified"
 /note="f379; This 379 aa ORF is 29 pct identical (33 gaps) to 364 residues of an approx. 400 aa protein BEXD_HAEIN SW: P2236"
 /codon_start=1
 /transl_table=1
 /product="putative function in exopolysaccharide production"
 /protein_id="AAC74068.1"
 /db_xref="GI:1787218"
 /translation="MKRKIEFSVTLAVLSLCTACTVPGONLSTSNKDIETLPDNOY LLDKMNIVPTPGLIDOLRAKPIMSQANPELEQIANVEYRIGIGDVLATVMDHPE LTPAGQVRSASDQNMVNDGALFYPIGKAKATLTVRMEITRRLDSVLESPO VDSVAFRSKATVTEGEVSKSGQPTINILITMDATNAAGSLTADMDRNVLTQ GVKTKVLYALMQGLRQNKLLHGDLEIPRNDLKVFVAGVGKOSTLKMDSGM TLAEALNABEMQDVAATGIFVIRATONQNGKIANTYOLNKKDASAMILGTEFOL EPRDIVYTTAPLARMNRVISLAVPTISGVHDLTETSRWIOIWTWN"
 complement(5355..7451)
 /gene="ymca"
 /note="b0984"
 complement(5355..7451)
 /gene="ymca"
 /function="orf: Unknown"
 /note="f698; This 698 aa ORF is 64 pct identical (7 gaps) to 694 residues of an approx. 704 aa protein YBH_ECOLI SM: P32689"
 /codon_start=1
 /transl_table=1
 /product="orf, hypothetical protein"
 /protein_id="AAC74069.1"
 /db_xref="GI:1787219"
 /translation="MKRSYLSCLATAVSSACHAEVLTYPDPGLSSQSDSGTGILLO MPNARIPEGESVNYRNDQYRFTSVSLAFPMLEGTIRITVDTRKYSQWEPDSGD QSTKDSFDEKLRLERSTYLPQVAFGRDLAGTGFDEGYLVASKQAGPDPFLGMA WGVANGNNTINPCRVSDKYCHRAESHADGDISFDFRGPASIFGIEYOTWNPPL

Alignment Scores:
 Pred. No.: 5,98e-163 Length: 10784
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 1 Gaps: 0

US-09-777-566a-2 (1-440) x AE000200 (1-10784)

QY	1	MeLysAlaIleuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla	20
DB	123	ATGAACGCGATCTAATCCCATTTTATCTCTCGATTCCGTTAACCCCGCAATCTCA	182
QY	21	PheAlaGlnSerGluProGluLeuLeuLysLeuSerValValIleValSerArgHisGly	40
DB	183	TTCCCTCAGAGTGGCCGAGAGCTGAAGCTGGAAGTGGTGAATGCTCAAGCGCATGCT	242
QY	41	ValArgAlaProThrIleValSerAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro	60
DB	243	GTCGGTCTCCACACCAAGCCACCACTGATGCAGATGTCACCCCGCAGCATGGCCA	302
QY	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuAlaTyrLeu	80
DB	303	ACCTGGCCGGTAAACAGCTGGTGGCTGCACCCGCGGTGTGAGCTAATCGCCATCTC	362
QY	81	GlyIleThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
DB	363	GGACATTACCAACCCCGAGCTGCTGATGCCAGCATGCTGGGGAAGGCGTGGCCG	422
QY	101	GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
DB	423	CAGTCTGCTCAGTGGCGGATTAATGCTGATCTGCAGACGAGCTACCCGTAACAGGCGAA	482

QY	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
DB	483	GCCTTCGGCGCGGCTGGACCTGATCTGCAATTAACCGTACATACCAGGACATACG	542
QY	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
DB	543	TCCAGTCCCGATCCGCTTATTTAATCTCTAAACATGGCGTTTGGCAACTGGATTAACCG	602
QY	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
DB	603	AACGTACTGACGCGATCTCAAGCAGGAGAGGAGGTCAATGCTGACTTACCGGCAT	662
QY	181	ArgGlnThrAlaPheArgGlyLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly	200
DB	663	CGGCAACGCGGTTTCCGGAACTGGAAGGGGTCTTAATTTCCGCAATCAACCTGTGC	722
QY	201	LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
DB	723	CTTAAACGCTGAGAAACAGCAGGAAAGCTTTCATTAAACGAGCGCTTACATCGCACTC	782
QY	221	LysValSerAlaAspAsnValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr	240
DB	783	AAGGTAGCGCGCAATGCTCTATTACCGGTGGTAAAGCTCGCATCATGCTGACG	842
QY	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr	260
DB	843	GAGATATTTCCTCTCAACAAGCAGCAGGAATGCCGAGCGGGGTGGGAAGATCACG	902
QY	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
DB	903	GATTCAACACAGATGGACACCTTGCTAAGTTGGCTTAAACGGCAATTTATTTTGGTACA	962
QY	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
DB	963	CGCAGCGCAGAGTGGCCCGCAGCGCGCCACCCCGTTATTAAGATTATCAAGACAGC	1022
QY	301	LeuThrProHisProProGlnGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
DB	1023	TTGACGCCCATCCACCCCAAAACAGCGCTAAGTGTGACATTAACCCACTTCAGTGG	1082
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnThr	340
DB	1083	TTTATCGCGGAGACGATCTAATCTGCAAAATCTGGCGGGCGCATGGAGCTCAACTGG	1142
QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGluArgThr	360
DB	1143	ACGCTTCCCGTCAAGCCGGATTAACAGCGCCAGGTGTGAACGTGTTTGAACGCTGG	1202
QY	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
DB	1203	CGTGGCTTAACGATTAACAGCCAGTGCATTCAGGTTTCGTCGCTTCCACACTTAAG	1262
QY	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr	400
DB	1263	CAGATGCGGTAAACCGCCCTGTCAATTAATACGCCCGCGGAGAGTGAATGTGCC	1322
QY	401	LeuAlaGlyCysGlyGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
DB	1323	CTGGCAGGATGTGAAGACGAAATGCGCAGGCAATGTGTTGGTGGCAGGTTTACGAA	1382
QY	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
DB	1383	ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG	1418
RESULT 14			
LOCUS	AX042378	17732 bp	DNA linear
DEFINITION	Sequence 7 from Patent WO0064247.		
ACCESSION	AX042378		
VERSION	AX042378.1	GI:11340996	
KEYWORDS	synthetic construct.		
SOURCE			

ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 17732)
 AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
 TITLE Transgenic animals expressing salivary proteins
 JOURNAL Patent: WO 0064247-A 7 02-NOV-2000;
 UNIVERSITY OF GUELPH (CA)
 FEATURES
 source 1. 17732
 location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32650"
 /note="Lama2/APPA transgene"
 BASE COUNT 4719 a 4125 c 4168 g 4719 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.08e-162 Length: 17732
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX042378 (1-17732)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuProLeuThrProGlnSerAla 20
 DB 12653 ATGAAGCCATCTTAATCCCATTTTATCTCTGATTCGCTTAACCCGCCAATCTGCA 12712
 QY 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 12713 TTGCTCAGAGTAGAGCCGAGTAGCAAGCTGAAGATGTGATTCAGTCGATGAT 12772
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 12773 GTGCGTGTCAACCAAGGCGACGCACTGATGAGATGTCACCCAGAGCGATGGCCA 12832
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgLysGlyGlnLeuIleAlaTyrLeu 80
 DB 12833 ACCGCGCGGTAAACGAGGTTGGTGTACACCGCGCGTGTGATGCTATACCTATCTC 12892
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 12893 GGCACATTAACCAAGCGCGCTGTGTGACCGAGGATGTCGGCGAAGAGGCTGCCCG 12952
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspLysArgThrArgLysThrGlyGlu 120
 DB 12953 CAGCTGTGTAGAGTGGGATTAATGCTGATGTGACAGCGGTACCCGTAAGACAGCGAA 13012
 QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 13013 GCCTTCCCGCCCGCGCGCTGACTGTGCAATACCTACTACCTACCGAGATACG 13072
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 13073 TCCAGTCCCGATCCGTTATTAATCTCTAAGAACTGGCGTTGCCAAGGATGAACGCG 13132
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
 DB 13133 AACGTGCTCAGCGAGTCCAGACAGGCGAGGAGGCAATTCCTGACTTACCGGGCAT 13192
 QY 181 ArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 13193 CGGCAAAAGCGCGTTCCGCGACTGGAACGGGTCTTAATTTTCCGCAATCAACTGTGTC 13252
 QY 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 DB 13253 CTAAACGTGAGAAACGAGCAAGAAAGCTGTTCAATTAACGAGCATTAACATCGAAGCTC 13312
 QY 221 LysValSerLysAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 13313 AAGGTGAGCGCGACATATGTCATTAACCGGTGGGTGAAGCCCTGCATCAATGCTGACG 13372

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr 260
 DB 13373 GAGATATTTCCTCCGCAACAACGACAGGAGATGCGGAGCGGGTGGGAGAGATCAC 13432
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnIlePheTyrLeuLeuGln 280
 DB 13433 GATTCACACAGTAGTGAACACCTTCCTAAGTTGATTAACGCGCAATTTTATTCCTCAAA 13492
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 13493 CGCAGCGCAGAGGTTGGCCGAGCGCGCCACCCGTTATTAGATTGATCAACAGACAGCG 13552
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 13553 TTGACGCCCATCCACCGCAAAAACAGCGGATGTGTGACATTAACCACTCACTGCTG 13612
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 13613 TTTATCCCGGACACGATTAATCTGCAAAATCTCGCGCGCGACACTGCACTCACTGG 13672
 QY 341 ThrLeuProGlnGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
 DB 13673 ACCTTCCCGGTACGCGGATTAACACCGCCGAGGTGTGAATGTGTTGAACGCTGG 13732
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 13733 CGTGGCTAAGCAATACAGCACTGATTCAGTTTCGCGCTTCACACTTACAG 13792
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
 DB 13793 CAGATGCGGTAAACGCGCTGTCTAATTAATACCGCCCGCGAGAGGTGAACCTGACC 13852
 QY 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGCAGAGTGTGAAGCGAAATGCCAGGCGCATGTGTTGTTGGAGGTTTAAACGAA 13912
 QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGTGAATGAAGACGATACCGCGTTGCGAGTTTG 13948
 RESULT 15
 D90735 18882 bp DNA linear BCT 07-FEB-1999
 LOCUS Escherichia coli genomic DNA. (22.3 - 22.7 mln).
 DEFINITION D90735 AB001340
 ACCESSION D90735.1 GI:4062539
 VERSION Complete and shotgun sequencing; hyaB; hyaC; hyaD; hyaE; hyaF;
 appC; cyxA; cbdA; appB; cyxB; cbdB; appA; amsI; insA; insB; cspF;
 cspB; sfa; yccL; yccM.
 SOURCE Escherichia coli (strain:K12) DNA, clone:Kohara clone #225.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 1 (sites)
 AUTHORS Oshima, T., Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A.,
 Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K.,
 Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T.,
 Mizobuchi, K., Mori, H., Motomura, K., Nakamura, I., Nishimoto, H.,
 Nishio, Y., Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K.,
 Wada, C., Yamamoto, Y., Yano, M. and Horiiuchi, T.
 TITLE A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 mln region on the linkage map
 JOURNAL DNA Res. 3 (3), 137-155 (1996)
 MEDLINE 97061202
 REFERENCE
 2 (sites)
 AUTHORS Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiiuchi, T.,
 Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K.,
 Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M.,
 Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H.,
 Motomura, K., Nakamura, I., Nishimoto, H., Nishio, Y., Oshima, T.,
 Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
 Yamamoto, Y. and Yano, M.
 TITLE The systematic sequencing of the Escherichia coli genome in Japan

JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 18882)
 AUTHORS Mori, H.
 TITLE Submitted
 JOURNAL Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.: 8916-5 Takayama, Ikoma, Nara 630-01, Japan (E-mail:hmori@egcc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)
 On Dec 24, 1998 this sequence version replaced gi:1651473.
 Collaboration Information:

COMMENT.
 Project:
 The Japan E.coli genome DNA sequencing project

Group:
 The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)
 Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.

Headed by:

Name: Takashi Horuchi
 Address: National Institute of Basic Biology, Okazaki, 444, Japan
 E-mail: kishori@nibb.ac.jp
 Information operator:
 Name: Hirotsada Mori
 Address: NARA Institute of Science and Technology,
 Ikoma, 630-01, Japan

E-mail: hmori@egcc.aist-nara.ac.jp
 URL:

The Japan E. coli genome database
 http://sw3.aist-nara.ac.jp.

FEATURES
 source

Location/Qualifiers

1. 18882

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="22.3-22.7 min"

/clone="Kohara clone #225"

/note="Nucleotide position 1034630-1053511 from the initiation site of *thra* (0 min.). This clone is from Kohara Lambda miniset library."

1. 966

/gene="hyab"

<1. 966

/note="ORF_ID:0224#6

similar to SwissProt Accession Number P19927"

/codon_start=1

/transl_table=1

/product="Hydrogenase-1 large chain (EC 1.18.99.1) (nife hydrogenase) (membrane-bound hydrogenase 1 large subunit)

(hyal) "

/protein_id="BA35738.1"

/db_xref="GI:4062540"

/translation="PDALAIQGFNKPMSSEITGSLDKCVLSYGAFFDIANDFGEKSL

MPGAGVINGDENNVLPVLDVPOQVEFVHAMRYRPNDOYGRHDFDITDPYNPBG

VKGSIDINQOLNEQERYSMIKAPRRGNAMEVGLPRLTILVHKGDATVTSVDRMS

ALNPLSGISQSTLGRITLCRAHEAQAAGKLOLFFDKLNTLNKGNLATASIEKEPAT

WPTKRGVGFTEAPRGALGHMAAIRDGIDIDYQCVPTTNASPPDPKGOIGAVEAAL

MNTMAIPEQPLRLTLHSPDLCLASHVLDGDSGLISVOVR"

985. 1692

/gene="hyac"

985. 1692

/gene="hyac"

/note="ORF_ID:0225#1

similar to PIR Accession Number JY0074"

/codon_start=1

/transl_table=1

/product="hyac protein"

/protein_id="BA35739.1"

/db_xref="GI:1651475"

/translation="MOQSDNVVSHYVEAPVRIMWLTVLCAVLAIVTGYFGKPLP

SVSGEATYLFMYGIRLIFSAGVETVLLMIRITMAVGRNSHELPIVPMKSM

OGWAEIRWYFLIARPSADIGHNIQAAGFVGLVSMVPIITSEFALYSESOAIF

APRRVVEFFYWTGNSMDIHSNRLGMLIGAFYIGVYALARDINDSDTIVSTM

NGYRSKRGKISNERS"

1689. 2276

/gene="hyad"

1689. 2276

/gene="hyad"

/note="ORF_ID:0225#2

similar to SwissProt Accession Number P19930"

/codon_start=1

/transl_table=1

/product="Hydrogenase-1 operon protein hyad."

/protein_id="BA35740.1"

/db_xref="GI:1651476"

/translation="MSQRYVYVYMGILNLMADDEGRVYAEKLYAHYHPEVEYVDG

GTQGLNLGYESASHLLILDAIDGLEPGLRTYAGERIPAYISAKTKMSLHNSFSE

VLALADIRGHLPAILALVGLQPMAMDYDGSLSLRLAQLPAEQALQAALAAWGI

OPANESRLNTDCSMENYEGVRLRYQWTOEEG"

2273. 2671

/gene="hyae"

2273. 2671

/gene="hyae"

/note="ORF_ID:0225#3

similar to SwissProt Accession Number P19931"

/codon_start=1

/transl_table=1

/product="Hydrogenase-1 operon protein hyae."

/protein_id="BA35741.1"

/db_xref="GI:1651477"

/translation="MSNDTPPDALMORMARGWTPVSESRDLDTQAPDGVLLSSD

PKRPEVSDNPMIGELRPEPDYQVAIADLQSESAIGDRFGVFRPATLVFTG

YRGVNGIHMPAELINLMRGLVEPQERAS"

2668. 3525

/gene="hyaf"

2668. 3525

/gene="hyaf"

/note="ORF_ID:0225#4

similar to SwissProt Accession Number P19932"

/codon_start=1

/transl_table=1

/product="Hydrogenase-1 operon protein hyaf."

/protein_id="BA35742.1"

/db_xref="GI:1651478"

/translation="MSEYFHLPGTQPNDSFPMNLPITCOVNDSPMAALEQCA

HSPQVIALNLELOHLSEROPPLGEVLVDLLNADNRHFNITLGGESVSRIOCA

DDSESEIOEAFICGLMWRVRRRGEKLEEDKLEACGAPALAAQATONLIPDLSLPP

IDGLNGLPLAHELIAHYRNPDPADPHSINTQLPISENDRLFLSLCGPGNITQITIG

YGESYINATYGRHWHLRCDTLKGPILLESYEICPIPEVILAAPDLVDSQORLSEVC

QMLAEAPPT"

3659. 5203

/gene="appc"

3659. 5203

/gene="appc"

/note="ORF_ID:0225#5

similar to SwissProt Accession Number P26459"

/codon_start=1

/transl_table=1

/product="Cytochrome bd-11 oxidase subunit 1 (EC 1.10.3.-)"

/protein_id="BA35743.1"

/db_xref="GI:1651479"

/translation="MMQVYIDLSRQFALTYHFLVPLTGLIFLILAMEIYVVG

KTIYDMTRFMWGLKPGINPAGVATGLTMEQFGNMSFYSNYVDIGDITQPLAMALM

AFLESTFVGLFFFGWRLNRYQHLVYVLAFAFGSNLSALMLNANGMGPPTGAHFD

IDTLMEVTSFSELVFNVSQVRFVYVYAGVYGAFTMAISAWYLLRGRERNVALR

SFALGSVFTGLAIGITQLDGDSAAVEVAOVQPVKLAAMEGMEQTEPAPAPHVVAALR

QDOERNAPALAKIPALGLITATWISLDPVPGKLNKLAETYPRLQGRMAMLMQETISOG

gene

CDS

NRBPHVLAQAFRLGEGDLYGMLLSRYAPDNHNVYAQYOAMRGALPOVAPEWMSFRI
 MVGSGSLILVLMIALVOTLRKIDQHRWLMKALMLPLPIALIEAGMPTERROP
 MAIODILPTYSANSAALTTGGLARSLIMIVGLYTLFLAEYLMQKVARLBSANQSRQ
 PTQOQC"

5215..6351
 /gene="appb"
 /note="cbbd; cyxb"
 5215..6351
 /gene="appb"
 /note="ORF_ID:0225#6
 similar to SwissProt Accession Number P26458"
 /codon_start=1
 /transl_table=1
 /product="Cytochrome bd-11 oxidase subunit 11 (EC
 1.10.3.-)"
 /protein_id="BA35744.1"
 /db_xref="GI:1651480"
 /translation="MEDYETLRPIWMLLIGVILVFMISDGFMDGICLLPLVARND
 ERRLVINSVAHMEQNVMLIAGALFAAMPRYAAAFSGFYVAAILVLSLEFRL
 AFYRGKISADARMKMDAGLVIGSLVPVFGAFNLLIGVFPATPOLRVEYLS
 FMOLLPEPLICGLSLGVNIILOGVMLOKTVGVHILRSOLATBRAALLVMCEFLA
 GYMLWVGDGFVLIADANGSPNIMKIVLIVPGAMNNNTVESPVMLIFPLGFCFL
 LTVMAIRGRPMGFLMASLMQRGVITFAGITLFPFVMPSSVSPISSLTIMDSTISOL

Alignment Scores:

Pred. No.: 1.16e-162 Length: 18882
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 1 Gaps: 0

US-09-777-566a-2 (1-440) x D90735 (1-18882)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 |||||
 DB 6536 ATGAACCGATCTTAATCCATTTTATCTCTCTGATTCGTTACCCGCCCAATCTGCA 6595
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValValIleValSerAlaHisGly 40
 |||||
 DB 6596 TTCCCTAGAGTGGCCGCGACCTGAGCTGGAAGAGTGGTGGATTTGTCAGTCGTCATGCT 6655
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 |||||
 DB 6656 GTGGGTGCTCCAAACGAGCCAGCCAACTGATGAGATGTCACCCAGAGCATGGGCA 6715
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 |||||
 DB 6716 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGGATGATGATGCTATCTC 6775
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 |||||
 DB 6776 GGACATTAACCAACCCGACGCTGCTGACCCGAGATTCGCTGGGAGAAAAGGCTGCCCG 6835
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 |||||
 DB 6836 CAGTCTGCTGAGTGGCGATATATGATGTCAGCAGCGCTACCCGTAAGAACAGCGCA 6895
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 |||||
 DB 6896 GCCTTGCCCGCGGGCTGGCCACTGACTGCAATACCGTACATACCCAGCAGATPACG 6955
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValLysGlnLeuAspAsnAla 160
 |||||
 DB 6956 TCCAGTCCCGATCCGTTATTTAATCTCTTAATAACTGGCGTTGCAACTGGATTAACGCG 7015
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 |||||
 DB 7016 AACGTGACTGACGGGATCTCAGCAGGCGCAGAGGATCAATTGCTGACTTACCGGCGCAT 7075
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
 |||||
 DB 7076 CGGCAACGCGGCTTCCGAGACTGGAACGGGCTGCTTAATTTCCGCAATCAAACTTGTC 7135

QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 |||||
 DB 7136 CTTAAACGTGAGAAACAGGAGCAAAAGCTGTTTCATTAACGACGATTAACCATGGAACCTC 7195
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 |||||
 DB 7196 AAGGTGAGCGCGCAGATGCTCATTAACCGGTGGGTAAAGCCTCGCATCAATGCTCAGC 7255
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 |||||
 DB 7256 GAGATATTTTCTCCGCAACAGCAGAGGATGCGGAGCGGGGTGGGAGAGATCAC 7315
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 |||||
 DB 7316 GATTCACACCAAGTGGAAACCTTCTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTCA 7375
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 |||||
 DB 7376 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCGTTAATAGATTGATCAAGACAGCG 7435
 QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 |||||
 DB 7436 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGACATTACCCACTTCAGTCTG 7495
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 |||||
 DB 7496 TTTATCCCGGACACGATACTAATCTGCAAAATCTCGGGCGGACACTGGAGACTCACTGG 7555
 QY 341 ThrLeuProGlnGlnProAspAsnThrProProGlyLysGluLeuValPheGluArgTrp 360
 |||||
 DB 7556 ACAGTCCCGGTACAGCGGATTAACACCGCCGACAGGTGTAACAGTGTGTTGAACGCTGG 7615
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 |||||
 DB 7616 CGTGGCTTAACGATTAACAGCAGCACTGATTCAGGTTTCGCTTCCTTCAGACTTTACAG 7675
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 |||||
 DB 7736 CTGGCAGAGTGTGAAGAGCGAAATGCCAGGCGCATGTGTTGCGCAGGTTTAAACGCAA 7795
 QY 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 |||||
 DB 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 |||||
 DB 7796 ATCGTGAATGAAGCAGCATACCGCGGTGCAGTTTG 7831

Search completed: June 12, 2003, 14:59:54

Job time : 3669 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 12:09:10 ; Search time 328 Seconds

(without alignments)
3020,973 Million cell updates/sec

Title: US-09-777-566a-2

Perfect score: 2302
Sequence: 1 MKAIIPLFLSLIPLFQSA.....IVNEARIPACSLRHHHHH 440

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV-xml
-O/cgnt2.1/USPRO.spool/US09777566/runat_05062003_154807_6525/app.query.fasta.1.583
-DB-n_Geneseq_101002 -QEXT-fastap -SUFFIX-rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCAITGN=200 -THR-SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT-pto -NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09777566 @CGN.1.1.263 @runat_05062003_154807_6525 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUERY -NR3_SCORES=0 -NAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	22 AAC88885	Escherichia coli B
2	2302	100.0	1323	24 AAC88885	Escherichia coli B
3	2302	100.0	1323	24 AAD25460	Escherichia coli B
4	2235	97.1	1299	21 ABK12514	DNA encoding phyt a
5	2235	97.1	1901	24 AAD25463	Escherichia coli a
6	2235	97.1	3470	21 AAC68298	R15/APPA plasmid c
7	2235	97.1	4060	21 AAC68296	R15/APPA plasmid c
8	2235	97.1	6116	21 AAC68297	R15/APPA plasmid c
9	2235	97.1	6708	21 AAC68295	R15/APPA plasmid c
10	2235	97.1	17732	21 AAC68300	Lama2/APPA plasmid
11	2235	97.1	20623	21 AAC68294	Lama2/APPA plasmid
12	2221	96.3	1489	21 AAA28216	E. coli acid pho
13	2217	96.3	1486	22 AAD06831	E. coli acid phosp
14	2205	95.8	5421	21 AAC68299	SV40/APPA plasmid
15	2197	95.4	1486	22 AAD06832	E. coli acid phosp
16	2194.5	95.3	1272	20 AAX26540	DNA encoding an Es
17	1401.5	60.9	846	22 AAS41940	Genomic sequence #
18	1196.5	52.0	1737	23 AAS89885	DNA encoding novel
19	813	35.3	519	22 AAS41139	CDNA encoding novel
20	759.5	33.0	4050	23 AAS64582	DNA encoding novel
21	657	28.5	1264	23 AAS88443	DNA encoding novel
22	657	28.5	1264	23 AAS94316	DNA encoding novel
23	539	23.4	11710	24 AHA43718	E. coli genome sec
24	378.5	16.4	356	23 AAS89883	DNA encoding novel
25	352	15.3	5975	15 AAO55758	Klebsiella pneumon
26	274	11.9	5976	23 AAS8450	DNA encoding novel
27	229	9.9	898	23 AAS75019	DNA encoding novel
28	195.5	8.5	1282	23 AAS87535	DNA encoding novel
29	195	8.5	1281	24 AAS17591	DNA encoding novel
30	195	8.5	1352	24 AABK14473	Human protein phos
31	184.5	8.0	3189	23 AAS82488	DNA encoding novel
32	184.5	8.0	3189	23 AAS89136	DNA encoding novel
33	166	7.2	427	21 AAA67371	Pinus radiata gluc
34	136.5	5.9	1272	24 ABL36298	Human lysosomal ac
35	133	5.8	1584	23 ABL11483	Drosophila melanog
36	124.5	5.4	1197	23 ABL29887	Aspergillus fumiga
37	119	5.2	1398	22 AAD17943	Drosophila melanog
38	119	5.2	1410	23 ABL09027	Aspergillus fumiga
39	119	5.2	1571	19 AAX03144	Aspergillus fumiga
40	119	5.2	1571	20 AAX27422	A. fumigatus phyt a
41	117	5.1	1588	18 AAT72721	PAP-GW-CSF immuno
42	117	5.1	3061	19 AAX29653	Human prostate can
43	117	5.1	3061	20 AAX26062	US5882864 Seq ID N
44	117	5.1	3061	21 AAZ87547	Prostatic acid pho
45	117	5.1	3061	22 AAS03766	Biomarker UC band

ALIGNMENTS

RESULT 1
AAC88885 standard; DNA; 1323 BP.

07-MAR-2001 (first entry)

Escherichia coli B phytase enzyme nucleotide sequence.

Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

nutrition; ds.

Escherichia coli.

WO200071728-A1.

30-NOV-2000.

XX 15-MAY-2001; 2001WO-US15764.
 PF
 XX 25-MAY-2000; 2000US-0580937.
 PR
 XX (DIVE-) DIVERSA CORP.
 PA
 XX Short JM, Kretz KA, O'Donoghue E;
 PI
 XX MPI: 2002-164149/721.
 DR
 XX P-PSDB: AAE22836.
 DR
 XX New dietary aids comprising sustained release biocompatible
 PT compositions, comprise agent that assists in digestion, useful for
 PT delivering enzymes, therapeutics, medicine or agents to an organism
 XX
 PS Disclosure; Fig 1; 89pp; English.
 XX
 CC The present invention relates to novel dietary aids comprising sustained
 CC release biocompatible composition which comprises an agent (enzymes such
 CC as phylase, amylase, esterase, protease) that assists in digestion. The
 CC biocompatible composition is effective upon oral consumption and release
 CC in the digestive tract of a subject. The dietary aids are useful for
 CC delivering enzymes, therapeutics, medicine and agents to an organism.
 CC The use of enzymes and other agents in digestive aids of livestock or
 CC domesticated animals not only improves the animal's health and life
 CC expectancy but also assists in increasing the health of livestock or
 CC in the production of foodstuffs from livestock. The present sequence
 CC is Escherichia coli phylase DNA.

XX SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 1 other;

Alignment Scores:

Pred. No.: 5.79e-202 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-777-566a-2 (1-440) x AAD36473 (1-1323)

QY 1 MetLysAlaIleuSerProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 DB 1 ATGAAGGATCTTAATCCATTTTATCTCTCGATTGCGTTAACCCCAATCTCA 60
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 61 TTCCCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTCGTCATGCT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGGGTGCTCCACCAAGCCAGCCAGCACTGATGCGAGTGTACCCCAAGCCAGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIleu 80
 DB 181 ACGGGCCGGAATAACTGCTGCTGCTGACACCCGCGNGTGGTGAAGCTAATCGCCATCTC 240
 QY 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 DB 241 GGACATATCCAAACCCAGCCGCTGCTGACCGAGATGTGGCAAAAAGGGGTGCGCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluValGlyThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGCTCAGGTGCGCATTAATGCTGATGTGCGAGCGCTACCCGTAACAGCGCAA 360
 QY 121 AlaPheAlaIleValAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCGCGCGCGCTGCGACCTGCTGCAATACCGTCATACCCAGCGCATGCG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGGATCGTATTATTAATCCCTTAATAAACTGCGCTTGGCAACTGGAATACCGC 480

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACTGACGGGATCTTACGACGAGGAGGATCAATCTGCTTACCGGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnProGlnSerAsnLeuGly 200
 DB 541 CGGCAAAAGCGGCTTCCGGAACCTGAACGAGGCTGCTTAATTTCCGCAATCAACTGTCG 600
 QY 201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAGCTGGAAGAACAGAGCAAGAAAGCTTTCATTAAACGAGGATTAACATCGAACTC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACCGCCGCAATGCTCTATTAAACGGTGGTGAACCTCGCATCAATCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyLysArgIleThr 260
 DB 721 GAGATATTCTCTGCAACAAGACAGAGGATGCCGCGGAGGAGGAGGATGAC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlLeuGln 280
 DB 781 GATTCACACACAGTGAACACCTTCTAGTTGATTAACGCGCAATTTATTCTCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCAGAGGTTGCCCGCAGCCGCGCCACCCGCTTATGATTTGATCATGCGACG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATCCACCCCAAAACAGCGCATGCTGATGATTAACCATTCAGTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
 DB 961 TTTATATGCGGACAGCATCTAATCTGCAATCTCGGCGCGCATGAGCTCAACTG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 DB 1021 ACGCTCCCGGTACAGCCGATTAACAGCCGCGCAGTGTGAACCTGTTTGAACGCTG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGTAAAGCATTAACAGCCAGAGTGTGAGTGTGCTGCTCCACATCTTACAG 1140
 QY 381 GlnMetLysAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 1141 CAGATGCGGTATTAACCCGCTGCTAATTAATACGCGCCGCGGAGGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGGATGTGAAGACGCAAAATGCGCAGGCGCATGTGCTTGCGAGTTTACGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 DB 1261 ATCGGATGAAGCAGCATACCGCGCTGCAAGTTGAGATGTCATCACCATCACCATCAC 1320

RESULT 3
 AAD25460
 ID AAD25460 standard; DNA; 1323 BP.
 AC AAD25460;
 XX 26-MAR-2002 (first entry)
 DE Escherichia coli B phylase DNA.
 XX Escherichia coli B phylase DNA.
 KW Bacterial phylase: K12 appa phylase: protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; diopulping;
 KW non-alcoholic drink; biobleaching; B phylase; ds.
 OS Escherichia coli.

XX Key Location/Qualifiers
 FH 1..1323
 FT /tag- a
 FT /product- "E. coli B phytase protein"
 XX MO200190333-A2.
 XX 29-NOV-2001.
 XX 24-MAY-2001: 2001MO-US17118.
 XX 25-MAY-2000: 2000US-0580515.
 XX (DIVE-) DIVERSA CORP.
 XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;
 PI WPI: 2002-083108/11.
 DR P-PSDB: AAE15806.
 XX New bacterial phytase for e.g. improving the nutritional value of
 PT phytase-containing feedstuffs and subsequently improving the growth
 PT performance of an organism that consumes it, or in treating animal
 PT digestive systems
 XX
 XX Claim 2: Fig 1: 170pp; English.
 XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
 CC K12 appa phytase. The enzyme has phytase activity and improved thermal
 CC tolerance when compared with wild-type phytase. It has improved protease
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytase-containing feedstuffs and subsequently
 CC improving the growth performance of an organism that consumes it. In
 CC treating animal digestive systems, in feed treatment processes and for
 CC in vitro purposes related to research, discovery and development. They
 CC are also used for generating recombinant digestive system life forms,
 CC for producing or manufacturing alcoholic and non-alcoholic drinks based
 CC on the use of moulds, grains and/or plants, in biopulping and bio-
 CC bleaching where a reduction in the use of environmentally harmful
 CC chemicals that are traditionally used in the pulp and paper industry
 CC is desired and in the reduction or possible elimination of the need
 CC for mineral supplements, enzymes or therapeutic drugs for animals
 CC from the daily feed thus increasing the amount calories and nutrients
 CC present in the feed. The present sequence is a DNA encoding E. coli
 CC B phytase protein.
 XX
 SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 1 other;
 Alignment Scores:
 Pred. No.: 5.79e-202 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-777-566a-2 (1-440) x AAD25460 (1-1323)
 QY 1 MetLysAlaIleuLeuProHeuSerLeuLeuProLeuThrProGlnSerAla 20
 DB 1 ATGAAAGCATCTTAATCCATTATCTCTTCGATTCGTTAAACCCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACGCGAGCTGAAGCTGGAAGTGTGATGTCTCATGTCTATGCT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 121 GTGCGTGTCCCAACAGGCGACGCAACTGATGATGATGATGATGATGATGATGATG 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80

DB 181 ACCTGCGCGGTAAACCTGGTTGGCTGACACCGCGNGGTGTGAGCTAATGCGCTATCTC 240
 QY 81 GlyHisTryGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 DB 241 GGACATTTACCAACGCGACGCTGTGGTACCGGAGATTGCTGGCAAAAAGGCGTGGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGTGACGCTGCGCATTTATGCTGATGTGACAGCGATCCGTTAAACAGCGCA 360
 QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTGCGCGCGGCGCTGGCACCCTGCAATGCAATACCGTACATACCGAGCAATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAAGCTGGGTTGGCACTGATTAACGG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACTGACGCGATCTCTCAGAGCGAGAGGCTCAATTGCTGACTTTACCGGCAAT 540
 QY 181 ArgGlnThrAlaPheArgGlnLeuGlnLysValLeuAsnPheProGlnSerAsnLys 200
 DB 541 CCGCAAAAGCGGCTTGGCACTGGAACGGGTGCTTAATTTCCCAATCAAACTTGTGC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAAACAGAGAGCAAAAGCTGTCTATTACGACGATTTACCATCGGACATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGAGGCGCGCAAAATGCTTCATTAAACGGTGGGTGAAGCTGCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr 260
 DB 721 GAGATTTTCTCTGTCACACACACAGGAAATCCGAGCGGCGGTGGGAAGATCAC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 781 GATTACACACCATGGAACACCTTGGTAAGTTGCTTAACCGCCCAATTTATTTGCTACAA 840
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CCGACGCGCAGAGTTCGCCGACGCGCGACCCGCTTATGTGATTTGATCAATGCGACGG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaLysValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCGCCCATCCACCGCAAAACAGCGGTATGGTGTACATTACCCACTTCAGTACG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 961 TTTATTTGCGGACAGATCTAATCTGCAAAATCTCGCGGCGCATGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
 DB 1021 ACGCTTCCCGCTCAGCGGATTAACAGCGCGCAGGTGTGAACGTGTGTTAAAGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CTTGCGCTAAGCGATTAACAGCCAGTGTGATTCAGGTTTCCTGCTTCAGACTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
 DB 1141 CAGATGCGGTGAATAAAGCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAATGAC 1200
 QY 401 IleuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGCGCAGATGTGAAGAGCAATGCGCAGGCAATGTTCTCTTGGCAGGTTTTCGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 DB 1261 ATCGATGAATGAACGACGATACCGGCGTGCAGTTGAGATCTCATCACATCACATCAC 1320

RESULT 4

ABK12514 standard; DNA; 1299 BP.

AC ABK12514;

05-JUN-2002 (first entry)

DNA encoding phytase associated protein.

Phytase; gene: ds.

Unidentified.

FT	Key	Location/Qualifiers
FT	CDS	1..1299
FT	/*tag=	a
FT	/product=	"Phytase associated protein"

KR9086028-A.

15-DEC-1999.

25-MAY-1998; 98KR-0018810.

25-MAY-1998; 98KR-0018810.

(WOJ-) WOJIN CO LTD.

Bae HD, Forceburgh CW, Goloben S, Cheng KJ;

WPI: 2000-645078/62.

P-PSDB; AAU77775.

Novel phytase gene, recombinant phytase and usage thereof -
 Claim 1; Fig 2; 10pp: Korean.

The invention relates to a novel phytase gene, a recombinant phytase gene and their uses. This sequence encodes a phytase associated protein, described in the invention.

SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 other;

Alignment Scores:

Pred. No.:	8,14e-196	Length:	1299
Score:	2235.00	Matches:	430
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	97.09%	Indels:	0
DB:	21	Gaps:	0

US-09-777-566a-2 (1-440) x ABK12514 (1-1299)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGGAGATCTTAATCCCATTTTATCTCTTCGATTCGTTAACCCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTGCACCGGAGCTGAAGCTGGAAATGTGGTATGTCTCATGTATGTT 120
 QY 41 ValAlaGlnAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 121 GTGCTGCTCCAAACCAAGGCCACGCACTGATGACGATGTCCACCCAGACGATGGCA 180
 QY 61 ThrTrpProValLysLeuGlnLysLeuThrProArgGlyGlyGlnLeuIleAlaThrLeu 80
 DB 181 ACCTGGCCGTAACCTGGGTTGGCTGACACCGCGGGTGGTGAAGCTAATCGCTATCTC 240
 QY 81 GlyHisLysArgLysGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyGlyCysPro 100
 |||||||

DB 241 GGACATTAACCAAGCCAGCTCTGCTAGCCAGGATGTCGGCGAAGGAGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGGTGCGATTAATGCTGATGTCGACAGCGGACCGGTAACAGGCGAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGGCGCGGGCTGGACCTGACGTGCAATTAACGCTAATACCGACGAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCCTTAATTAATCTTAATAAAGCTGGCTGGCACTGGATTAACCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgLysGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AAGGTGACGACGCGATCCTCAGCAGGAGGAGGAGGTCATATGCTGACTTAACCGGAT 540
 QY 181 ArgGlnThrAlaPheArgGlnLeuGlnLysValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAAGCGCGCTTCGGGAAGCTGAGACGGGTCCTTAATTTCCGCAATCAACTGTGCG 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTGCTTCAATTAACGACGATTAACCGGACTC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyValAspSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACGCGCGCAATGTCATTAACGGTGGGTAAGCTGCAATGCTGACG 720
 QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyLysArgLysThr 260
 DB 721 CAGATATTTCTCTCCCAACCAACAGGAAAGCCGAGCCGGGTGGGAAAGATCAAC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGln 280
 DB 781 GATTCAACACAGTGGAAACACCTTGTGCTAAGTTGCAATACGGCAATTTTATTTGCTACA 840
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCAGAGGTTGCCCGCAGCGCCGCGCACTTATTAAGATTGATTAAGACACGCG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaLysGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTGTGACATTAACCACTCACTGCTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 961 TTTATCGCCGCGACAGATTAATCTGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
 DB 1021 ACGTTCGCCGTCACCGGATTAACACGCGCGGAGGTGTGAAGCTGTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTAAGGATTAACACGCAATGATTCAGCTTGGCTTCCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 1141 CAGATGCGGATTAACAGCGCGCTGCAATTAATACGCGCGCGGAGAGGTGAAGTACAC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnLysMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGAGTGTGAAGACGGAATGCGAGGCAATGCTGTTGGCAGGTTTATACGAA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 1261 ATCGTAATGAAGCACGCAATACCGGCTGCAAGTTTG 1296
 RESULT 5
 AAD25463
 ID AAD25463 standard; DNA; 1901 BP.

XX	AAD25463;
AC	
XX	26-MAR-2002 (first entry)
DT	
XX	
DE	Escherichia coli appa phytase wild type DNA.
XX	
KM	Bacterial phytase; K12 appa phytase; protease stability; anabolic;
KW	gastrointestinal; nutritional value; feed treatment process; therapy;
KM	thermal tolerance; growth performance; alcoholic drink; biopulping;
KW	non-alcoholic drink; biobleaching; ds.
XX	
OS	Escherichia coli.
XX	
FH	Key Location/Qualifiers
FT	CDS 188..1486
FT	/tag= a
FT	/product= "E. coli appa phytase protein"
PN	WO200190333-A2.
XX	
PD	29-NOV-2001.
XX	
PF	24-MAY-2001; 2001WO-US17118.
XX	
PR	25-MAY-2000; 2000US-0580515.
XX	
PA	(DIVE-) DIVERSA CORP.
PI	Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;
DR	WPI: 2002-083108/11.
XX	P-PsDB; AAEI5807.
PT	New bacterial phytase for e.g. improving the nutritional value of
PT	phytate-containing foodstuffs and subsequently improving the growth
PT	performance of an organism that consumes it, or in treating animal
PT	digestive systems
XX	
XX	Claim 42; Fig 7; 170pp; English.
PS	
CC	The patent discloses recombinant bacterial phytase from Escherichia coli
CC	K12 appa phytase. The enzyme has phytase activity and improved thermal
CC	tolerance when compared with wild-type phytase. It has improved protease
CC	stability at low pH. The recombinant phytase is useful for improving the
CC	nutritional value of phytate-containing foodstuffs and subsequently
CC	improving the growth performance of an organism that consumes it, in
CC	treating animal digestive systems, in feed treatment processes and for
CC	in vitro purposes related to research, discovery and development. They
CC	are also used for generating recombinant digestive system like forms,
CC	for producing or manufacturing alcoholic and non-alcoholic drinks based
CC	on the use of moulds, grains and/or plants, in biopulping and bio-
CC	bleaching where a reduction in the use of environmentally harmful
CC	chemicals that are traditionally used in the pulp and paper industry
CC	is desired and in the reduction or possible elimination of the need
CC	for mineral supplements, enzymes or therapeutic drugs for animals
CC	from the daily feed thus increasing the amount calories and nutrients
CC	present in the feed. The present sequence is a DNA encoding E. coli
CC	appa phytase wild type protein.
XX	
SO	Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 1 other;
Alignment Scores:	
Pred. NO.:	1,38e-195 Length: 1901
Score:	2235.00 Matches: 430
Percent Similarity:	99.54% Conservative: 0
Best Local Similarity:	99.54% Mismatches: 2
Query Match:	97.09% Indels: 0
DB:	24 Gaps: 0
US-09-777-566A-2 (1-440) x AAD25463 (1-1901)	

Db	188	ATGAAAGCATCTTAATCCATTTTTATCTCTTGATTCGGTTAAACCCGGCATCTGCA	247
Qy	21	phealaglnserglnuProgluleuLysleuGlnSerValValIleValSerArgHisgLy	40
Db	248	TTGCGTCAGAGTAGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGTATGTGCATCTGCATGGT	307
Qy	41	ValAlaGalaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
Db	308	GTGCGGTCTCCAAACCAAGGACCAAGCAACTGATCTCAGATGTCTCAACCCAGCGCATGGCCA	367
Qy	61	ThrTrpProValLysleuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTyLeu	80
Db	368	ACCTGGCCGGTAAACTGGGTTTGGCTGACACCGCGAGTGGTGAAGCTAAATGCCCTATCTC	427
Qy	81	GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro	100
Db	428	GGACATTTACCAACGCCAGCGTGTGTAAGCGACGGAATTGCTGGCGAAAAAGGCGTGGCCG	487
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln	120
Db	488	CAGTCTGGTCAAGTCCGATTATTTGCTGATGTGCAGAGGCTCAACCGTAAACAGGGAA	547
Qy	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleGlyThrValHisThrGlnAlaAspThr	140
Db	548	GCGTTGCGCCGGCGGGCTGGCACCTGACTGTGCATTAACGCTACATATCCAGCGAGATACG	607
Qy	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
Db	608	TCCAGTCCCGAGTCCGTTATTTAATCTCTAAAAACTGGCGTTGGCCAACTGATGAACGGC	667
Qy	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
Db	668	AACGTGACGTAGACGCGATCTCAACGACGAGGAGGCTCAATTCTGCTTACCGGGCAT	727
Qy	181	ArgGlnThrAlaPheArgGlnLeuGlnIleGlyValLeuAsnPheProGlnSerAsnLeuCys	200
Db	728	CGGCAAAACGGCGTTTCCGCAACTGGAAACGGGTCTTAATTTTCCGCATATAAACCCTGTGC	787
Qy	201	LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu	220
Db	788	CTTAAACGAGAGAAACAGACAGCAAGACTGTCTATTAAACGAGGCATTACATCGGAATCTC	847
Qy	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	848	AAGGTGAGGCGCGACAAATGCTCATTAACCGGTGCGGTAAAGCTCCATCAAAAGCTGACG	907
Qy	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr	260
Db	908	GAGATATTTTCTCTCGAACAAAGCACAGGGAATCCGGAACCGGGGTGGGAAAGATCAACC	967
Qy	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlLeuGln	280
Db	968	GATTCACACACAGTGGGAACACCTTGTCGTAAGTTTGCATAAACGCGCAATTTATTTGCTACAA	1020
Qy	281	ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	1028	CGCAGCGCCAGAGGTGGCCGCGACGCGCCACCCCTGATTAGATTGTGATCAAGACAGCGC	1080
Qy	301	LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu	320
Db	1088	TTTGACGCCCCCAATCCACGAAAAACAGGGGTATGGGTACACATTACCCACTTCAGTGCCTG	1144
Qy	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp	340
Db	1148	TTTATTCGCGGACAGATCTAATCTGGGCAAACTCTGGGCGGCACCTGGAGGCTCAACTGG	1200
Qy	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp	360
Db	1208	ACGCTTCCCGGTCAACGGGATTAACACGCCGCGCAGTGGTGAAGTGGTGTGTAACGCTGG	1260
Qy	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380


```

OY 381 GlnneLcAGApLyStHrProLeuSerLeuAnThrProPcOGlyGLValLySLeuThr 400
DB 2951 CAGATGCGTGAATAAACCCCGCTGTCATTAAATACGCCGCCCGGAGAGTGAACACTGACC 3010
OY 401 LeuAlaGlyCySGluGluGluGAsnAlaGlnGlyMeCySSerLeuAlaGlyPheThrGln 420
DB 3011 CTGGCAGGATGTGAAGACCGCAAAATGCGCAGGCGCATGTGTGCTTGCGCAGGTTTACGCCAA 3070
OY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB 3071 ATCGTGAATGAAGCACGCATACCCGCTTGCAAGTTTG 3106
RESULT 7
AAC68296
ID AAC68296 standard; DNA; 4060 BP.
XX
AC AC AAC68296;
XX
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid coding sequence.
XX
KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX environmental pollution; pig; ds.
XX
OS Rattus sp - chimeric.
XX
OS Escherichia coli - chimeric.
XX
PN WO200064247-AI.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000MO-CA00430.
XX
PR 23-APR-1999; 99US-0130508.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI Forsberg CW, Golovan S, Phillips JP;
XX
DR WPI: 2000-687245/67.
XX
DR P-PSDB; AAB36259.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PI expression of a protein, preferably phytase, comprises a nucleic acid
PI sequence including a heterologous transgene construct encoding the
PI protein -
XX
XX
XX Claim 14; Fig 19; 152pp; English.
XX
PS
XX
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case, pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
XX
SQ Sequence 4060 BP; 1257 A; 814 C; 843 G; 1146 T; 0 other;
Alignment Scores:
Pred. No.: 3,93e-195 Length: 4060
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 21 Gaps: 0
US-09-777-566A-2 (1-440) x AAC68296 (1-4060)
OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 1811 ATGAAGCCATCTTATATCCATTATTATCTCTTCGATTCCGTTAAACCCGCAATCTGCA 1870

```

QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
 DB 2951 CAGATGGCTGATAAAGCGCGCTGTCTATTAATACGCCGCCGAGAGGTGAACCTGACC 3010
 QY 401 LeuAlaGlySerGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 3011 CTGGCAGAGTGTGAAGAGCAAAATGCGCAGGCGCATGTGCTTGCGCAGGTTTACGCCAA 3070
 QY 421 ILeValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 3071 ATCGTGAATGACGACCATACCCGCTTGCAAGTTTG 3106

RESULT 8

AAC68297 standard; DNA: 6116 BP.

XX AAC68297;

XX 20-FEB-2001 (first entry)

XX R15/APPA plasmid coding sequence.

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 XX environmental pollution; pig; ds.

XX Rattus sp - chimeric.

XX Escherichia coli - chimeric.

XX W0200064247-A1.

XX 02-NOV-2000.

XX 20-APR-2000; 2000WO-CA00430.

XX 23-APR-1999; 99US-0130508.

XX (UYGU-) UNIV GUELPH.

XX Forsberg CW, Golovan S, Phillips JP;

XX WPI; 2000-687245/67.

XX P-PSDB; AAB36260.

XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein

XX Claim 56; Fig 20; 152pp; English.

XX The present invention provides transgenic animals which produce desired
 CC proteins. In this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.

XX Sequence 6116 BP; 1724 A; 1386 C; 1407 G; 1599 T; 0 other;

XX Alignment Scores:

Pred. No.: 6,32e-195 Length: 6116
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best local Similarity: 97.09% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 21 Gaps: 0

US-09-777-566a-2 (1-440) x AAC68297 (1-6116)

QY 1 MetLysAlaIleuIlePheProIleuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1811 ATGAAGCCATCTTAATCCCATTTTATCTCTGTGATCCGTTAACCCCGCAATCTGCA 1870

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 1871 TTGCTCAGAGTAGCGCGAGCTGACCTGGAAGTGTGGTGAATTCAGTCCATAGCT 1930
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 1931 GTGGCTGCTCCAAACCAAGGCCACCACTGATGAGATGACACCCAGACGATGGCCA 1990
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrLeu 80
 DB 1991 ACCTGCCCGGTAAGACTGGGTGGCTGACACCGCGCGGTGGTGAATTCGCTATCTC 2050
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspLysLeuLeuAlaLysGlyCysPro 100
 DB 2051 GGACATTACCAACCCACCGCTGTGTGACCGCAGGATTCGTGGCGAAAAGGCTGCCCG 2110
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 2111 CAGTCTGTCAGTGCAGGATTAATGCTGATGTCAGCAGCGCTACCCGTAAGACAGCGAA 2170
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 2171 GCTTTCGCGCGGGGCTGGCACCTGACTGTGCATTAACCTACCTACCCAGCGAGTACG 2230
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 2231 TCCAGTCCCGCATCCGTTATTTAATCTCTTAAAACTGGCGTTTCCCAATCGATTAACGGC 2290
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
 DB 2291 AACGTGACTGACGCGAGATCTCAGCAGGAGGAGGTCAATTCGCTGACTTACCGGGCAT 2350
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 2351 CGGCAAAAGCGGTTTCGGAACCTGAAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 2410
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 DB 2411 CTTAAAGCTGAGAAACAGAGCAAGAAAGCTGTTCAATTAACGACGATTAACCTGGAACCTC 2470
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 2471 AAGGTGAGCGCGCAATGCTCTCATTAACCGGTGTAAGCCCTGCATCAATCTGTACG 2530
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 DB 2531 GAGATATTTCTCTCGCAACAGCAGAGGAATGCCGAGCGGGGTGGGAAGATCAC 2590
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 DB 2591 GATTCACACACAGCTGGAACACCTTCTTAAGTTTGCATTAACCGCAATTTTATTCCTACAA 2650
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspIleIleMetAlaAla 300
 DB 2651 CGCACGCGCAGAGGTGGCCCGCAGCGCGCCACCCGTTAATGATGATCAACAGACGCG 2710
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 2711 TTGAGCCGCCATTCACCGCAAAACAGCGGATGTGTGATGATTAACCATCTCAGTCTGT 2770
 QY 321 PheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 DB 2771 TTTATCCCGCGACACGATACATTAATCTGGCAAAATCTCGCGCGCAGCGAGCTCAACTGG 2830
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 DB 2831 ACGCTTCGCGTACGCGGATTAACACCGCCGCAAGTGTGTAACGTGTTGAACGCTGG 2890
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 2891 CGTGGCTGAAGCGATTAACAGCAGTGAATTCAGGTTCGCGCTTCCAGACTTACAG 2950
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

DB 2951 CAGATGCGTATTAACCCCGCTGCATTAAATACGCCGCCGAGAGGTGAACCTGACC 3010
OY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB 3011 CTGCAGAGATGTGAAGACGAAATGCGCAGGCGCATGTGTGTCGAGGTTTTCGCA 3070
OY 421 IleValAsnGluAlaArgIleProAlaGlySerLeu 432
DB 3071 ATCGTAATGAAGCAGCATACCCGCTGTCAGATTGG 3106
RESULT 9
AAC68295
ID AAC68295 standard; DNA: 6708 BP.
XX AAC68295;
AC AAC68295;
XX 20-FEB-2001 (first entry)
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid coding sequence.
XX
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM environmental pollution; pig; ds.
XX
XX Rattus sp - chimeric.
OS Escherichia coli - chimeric.
XX
XX WO200064247-A1.
XX
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-CA00430.
XX
XX 23-APR-1999; 99US-0130508.
XX
XX (UNIV-) UNIV GUELPH.
XX
XX Forsberg CW, Golovan S, Phillips JP;
PI
XX WPI: 2000-687245/67.
DR P-PSDB: AAB36258.
XX
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein -
XX
XX Claim 56; Fig 18; 152pp; English.
XX
XX PS The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
XX Sequence 6708 BP: 1916 A; 1479 C; 1515 G; 1798 T; 0 other:
SO
Alignment Scores:
Pred. No.: 7,86e-195 Length: 6708
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 21 Gaps: 0
US-09-777-566A-2 (1-440) x AAC68295 (1-6708)
OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 1811 ATGAAAGCCATCTTAATCCCAATTTTATCTTCTTGATTCGTTAAACCCCAATCTGCA 1870
OY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40

DB 1871 TTCGTCAGAGTGAACCCGAGACTGAAGCTGGAAGTGGTGATTTGTCAGTCGATACGT 1930
OY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 1931 GTGGCTGCTCAACCAAGCCAGCAGCACTGATGAGATGTCACCCCGAGCGATGGCCA 1990
OY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
DB 1991 AACTGGCCGGTAAACCTGGTTGGTGGTGCACCCGGGGGTGGTGAAGTAAATGCGCTATCTC 2050
OY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB 2051 GGACATTTACCAACGACGACGCTGTGTGACCGAGATTGCGGGGAAAAAGGGTGGCCG 2110
OY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
DB 2111 CAGTCGTGTCAGGTCGCGCATTTATGCTGATGTCGACGAGCGTACCCGTTAAACAGGCGCA 2170
OY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 2171 GCTTCGCGCGCGGGCTGGACCTGACTGTGCATTAACCGTACATCCAGGAGATACG 2230
OY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 2231 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAAGTGGCGTTGGCACTGGATTAACGG 2290
OY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 2291 AACGTAGCTGACGCGCATCTCAGAGGCGCAGAGGCGTCAATTCGACTTTAACGGGCGAT 2350
OY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuGly 200
DB 2351 CGGCAACGGCGTTTGGCAACTGGAACGGGGCTTAATTTCCGCAATCAACTGTGGC 2410
OY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 2411 CTTAAAGCTGAAGAAACAGAGCAAGAAAGCTTTCATTAAACGAGCATTTACATCGGAATC 2470
OY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 2471 AAGGTGACGCGCGCAATGTCTCATTAAACCGGTGGTGAACCTCGCATCAATGTGACG 2530
OY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyLysArgIleThr 260
DB 2531 GAGATATTCTCTCCAAACACACACAGGAAATGCCGCGGGGTGGGGAAGATCC 2590
OY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
DB 2591 GATTACACACAGTGAACACCTGCTAGTTGCTTAACGCGCAATTTTATTTGCTACAA 2650
OY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspPheIleMetAlaAla 300
DB 2651 CGCAGCGCAGAGGTGGCCGCGAGCGCGCACCCCGTTATTAAGTTATCATCAAGACACCG 2710
OY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
DB 2711 TTGAGCCCATCCACCCCAAAACAGCGCATGTGCTGATGATTAACCATCTCACTGCTG 2770
OY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
DB 2771 TTTATCGCGGACACAGATACATTAATCTGCAATATCGCGCGGAGCTGAGACTCAACTCG 2830
OY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB 2831 ACGCTTCCCGGTCAACCGCATTAACCGCGCGCAGGTGTGAACGTGTTGAACGCGTGG 2890
OY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB 2891 CGTGGCTTAACGATTAACAGCCAGAGGATTAAGGTTTGCTGGTCTTCAACATTTTACAG 2950
OY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400

```

Db      2951 CAGATGCGTGATAAAGCCGCTGTCATTAATACGCCGCCGAGAGGTGAACCTGACC 3010
QY      401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
      |||
Db      3011 CTGGCAGAGATGTGAAGAGCCGAAATGCCAGGGCATGTGTGTTGTTGCGCAGTTTACGCA 3070
QY      421  IleValAsnGluAlaArgIleProAlaCysSerLeu 432
      |||
Db      3071 ATCGTAATGAAGCAGCATACCCGCTGCAAGTTTG 3106
      |||
RESULT 10
AAC68300
ID      AAC68300 standard; DNA; 17732 BP.
XX
AC      AAC68300;
XX
DE      20-FEB-2001 (first entry)
XX
DE      Lama2/APPA plasmid coding sequence.
XX
KM      Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX      environmental pollution; pig; ds.
OS      Mus musculus - chimeric.
OS      Escherichia coli - chimeric.
XX
PN      WO200064247-A1.
XX
PD      02-NOV-2000.
XX
PF      20-APR-2000; 2000WO-CA00430.
XX
PR      23-APR-1999; 99US-0130508.
XX
PA      (UYGU-) UNIV GUELPH.
PI      Forsberg CW, Golovan S, Phillips JP;
XX
DR      WPI: 2000-687245/67.
DR      P-PSDB; AAB36263.
XX
PT      Transgenic non-human animal for gastrointestinal tract specific
PT      expression of a protein, preferably phytase, comprises a nucleic acid
PT      sequence including a heterologous transgene construct encoding the
PT      protein.
XX
XX
PS      Claim 14; Fig 23; 152pp; English.
XX
CC      The present invention provides transgenic animals which produce desired
CC      proteins, in this case pigs which expresses phytase in the salivary
CC      gland. Low phytase production levels result in phytate in the diet being
CC      excreted and causing phosphorus contamination in water, as well as
CC      reducing the growth of animals. The invention provides a number of
CC      transgenes containing the E. coli APPA phytase coding sequence.
XX
SQ      Sequence 17732 BP; 4719 A; 4125 C; 4168 G; 4719 T; 1 other;

Alignment Scores:
Pred. No.:      3.01e-194      Length:      17732
Score:          2235.00      Matches:      430
Percent Similarity: 99.54%      Conservative: 0
Best Local Similarity: 99.54%      Mismatches: 2
Query Match:    97.09%      Indels:      0
DB:            21      Gaps:      0

US-09-777-566a-2 (1-440) x AAC68300 (1-17732)
QY      1  MetLysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
      |||
Db      12653 ATGAAGCCATCTTAATCCCATTTTATCTTCTGTGATCGGTTAACCAGCCCATCTGCA 12712
QY      21  PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
      |||

Db      12713 TTCCGTCAGAGTGAAGCCGAGACCTGAGAAAGTGTGATGTCAGTGCATAGT 12772
QY      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
      |||
Db      12773 GTGGCTGCTCCACACCAAGCCACCACTGATGCAAGATGTCACCCACAGACGTGGCCA 12832
QY      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
      |||
Db      12833 ACCTGGCCGGTAAACATAGGCTGTGGCTGACACCGCGGTGTGAGCTAATGCCATCTC 12892
QY      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspIleLeuAlaLysGlyCysPro 100
      |||
Db      12893 GGACATTACCAACCCAGCGCTGTGTGCCAGCATGTCGGGAAAGGGCTGCCCG 12952
QY      GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
      |||
Db      12953 CAGTGTGTCAGGTGCGATTATGTGATGTCACGAGCGGTACCGGTAACAGGGGAA 13012
QY      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
      |||
Db      13013 GCCTTCGGCGCGGGGCTGGCACCTGACCTGTCATTAACCGTACATACCCAGCGCATACG 13072
QY      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
      |||
Db      13073 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGGCTTCCCAACTGATAACGCG 13132
QY      161  AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
      |||
Db      13133 AACGTGACGACGCGATCTCAGCAGGCGCAGAGGCTCAATGCTGACTTACCGGGCAT 13192
QY      181  ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
      |||
Db      13193 CGGCAACGGCGTTTCGCGAATCGAAGCGGTGCTTATTTCCGCATCAATCAATTGTGC 13252
QY      201  LeuLysArgGluLysGlnAspLusCysSerLeuThrGlnAlaLeuProSerGluLeu 220
      |||
Db      13253 CTTAAGCTGAGAAACAGAGAACGCTTCAATTAACGCGCATTAACCATCGGAACCTC 13312
QY      221  LysValSerAlaAspAsnValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
      |||
Db      13313 AAGGTGACCGCGCAATGTCTTACCGGTGCGGTACCGTCCGCAATCAATCTGACG 13372
QY      241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
      |||
Db      13373 GAGATATTTCCTCGCAACAAGCAGCAGGAATGCCGAGCGGGGTGGGGAAGATCACCC 13432
QY      261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
      |||
Db      13433 GATTCAACACCCAGTGGAAACACCTTCTTAAGTTGCAATTAACGCGCAATTTATTTCTCA 13492
QY      281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
      |||
Db      13493 CGCACGCCAGAGGTTGCCCGGACGCGGCCACCCCGTTATTAAGATTGATCAACACAGCG 13552
QY      301  LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
      |||
Db      13553 TTGAAGCCCATTCACACGCAAAACAGCGCATGTGTGACATTAACCACTCAGTGCCTG 13612
QY      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
      |||
Db      13613 TTTATCCCGCGACACGATACCTAATCTGGCAATCTCGCGCGCACGAGCTCAACTGAG 13672
QY      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
      |||
Db      13673 ACGCTTCCGGGTCAAGCGGATTAACACGCGCGCAGGTGTGAATGCTGTGAACGCTGG 13732
QY      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnAlaSerLeuValPheGlnThrLeuGln 380
      |||
Db      13733 CGTGGCGTAAGCGAATACAGCAGCACTGATTAAGTTTCGCGTCCGACACTTACAG 13792
QY      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
      |||
Db      13793 CAGATGCGTGATAAAGCGCCGCTGCTCATTAATACGCGCGCGGAGAGAGTGAACCTGACC 13852

```

QY 401 LeuAlaGlyCysGluArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGAGAGATGTGAAGAGCAGAAATGCGACGAGCATGTGTTGGTTCAGGATTTACGCA 13912
 QY 421 IleValaGlnAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGAATGAGACGACGATACCGCTTGACAGTTTG 13948
 RESULT 11
 AAC68294
 ID AAC68294 standard: DNA; 20623 BP.
 XX AAC68294;
 AC AAC68294;
 XX 20-FEB-2001 (first entry)
 DT
 XX
 DE Lama2/APPA plasmid coding sequence.
 XX
 KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KM environmental pollution; pig; ds.
 XX
 OS Mus musculus - chimeric.
 OS Escherichia coli - chimeric.
 XX
 PN WO200064247-A1.
 PD 02-NOV-2000.
 PE 20-APR-2000; 2000WO-CA00430.
 PR 23-APR-1999; 99US-0130508.
 PA (UYGU-) UNIV GUELPH.
 PI Forsberg CM, Golovan S, Phillips JP;
 XX
 DR WPI: 2000-687245/67.
 DR P-PSDB: AAB36257.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein -
 XX
 PS Claim 56; Fig 5; 152pp; English.
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.
 XX
 SQ Sequence 20623 BP; 5449 A; 4847 C; 4902 G; 5424 T; 1 other;
 Alignment Scores:
 Pred. No.: 3,71e-194 Length: 20623
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservatave: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 21 Gaps: 0
 US-09-777-566a-2 (1-440) x AAC68294 (1-20623)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 12653 ATGAAGGCATCTTAATCCATTTTATCTTCGATTCGTTAAACCCCAATCTCA 12712
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 12713 TTCGCTCAGATGAGCCGAGACTGAGCTGGAAGAGTGAGATTGTCAGTGTCATGCT 12772

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 12773 GTGGTGTCTCCAAACCAAGCCACGCACTGATGAGATGTCACCCAGACGATGGCCA 12832
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 DB 12833 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGGTGGTGACTTAATGCGCTATCTC 12892
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 DB 12893 GGACATTACCAACGCCACGCTCTGTTACCCAGCATTCGTCGGCAAAAAGCGTCCCGC 12952
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 12953 CAGTGTGTCAAGGTGGCAGATTATTGCTGATGTCGACGAGCGTACCCGTAACAGCGCAA 13012
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 13013 GCCTTCGCCCGCGGCTGGCACCTGACTGCTCAATAACCGTACATACCAGCAGATPAG 13072
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 13073 TTCAGTCCCGATCCGTTTATTAATCTTAAAACTGCGCTTGGCAACTGATPACGCG 13132
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 13133 AACGTACTGACGCGATCTCAGCAGGCGAGAGGTCATATGCTTACCGGCGCAT 13192
 QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 13193 CGGCAAACGGCGTTCGCGCACTGGAAAGGGGCTTAATTTCCCAATCAAACTGTGTC 13252
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 13253 CTTAAACGTGAGAAACAGCAGCAAGCTGTTCAATTAACGACGATTAACATCGGAATTC 13312
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 13313 AAGGTAGCGCGCGCAATGCTCATTAACCGGTGCGTAAACCTCGCATCAATGCTGACG 13372
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlyTrpGlyArgIleThr 260
 DB 13373 GAGATATTTCTCTGCAACAGCACAGGGAATGCGGAGCGGGGTGGGAAGGTCCACC 13432
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 13433 GATTACACCAAGTGAACACCTTCTAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 13492
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 13493 CGCAGCGCAGAGGTGCGCGCAGCGCGCCACCCCGTATTAAGATTATCATCAAGACACG 13552
 QY 301 LeuThrProHisProProGlnLysGlnAlaArgValAlaThrLeuProThrSerValLeu 320
 DB 13553 TTGACGCCCATCCACCCCAAAAACAGCGTATGTTGATGATTAACCACTTCAAGTGTG 13612
 QY 321 PheIleAlaGlyHisAspThrAsnLeuValAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 DB 13613 TTTATCGCGGAGACAGATACATCTGCAATCTCGCGGCGGACGTGAGACTCAACTG 13672
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 DB 13673 ACGCTTCCCGGTCAACCGCATTAACCGCGCGCAGGTGTGAATGCTTGTGAACGCGTGG 13732
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 13733 CGTGGCTTAACCGAATACAGCAGGTGATTAAGTTTGTGCTTCCACACTTATACAG 13792
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
 DB 13793 CAGATGCGTGAATAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGAGTAAACTGACC 13852

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGCAGAGATGTAAGAGCAATGCGCAGGCATGTGCTGGCAGGTTTACGCA 13912
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGTGAATGACACGACATACCCGCTTCGAGTTTG 13948
 RESULT 12
 AAA28216
 ID AAA28216 standard; DNA: 1489 BP.
 AC AAA28216;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE E. coli acidic phosphatase appa2 encoding nucleotide sequence.
 XX
 KM Phosphatase; phytase activity; protease; acidic phosphatase; appa2;
 KM periplasmic phosphonhydrolase phosphonhydrolase; animal feed; ds.
 XX
 OS Escherichia coli.
 XX
 FH Key location/Qualifiers
 FH protein_bind complement (1..22)
 FT /tag- a
 FT /bound_moiety- "Primer Pfl"
 FT 16..108
 FT CDS
 FT /tag- b
 FT /product- "Appa2 related peptide"
 FT 184..1482
 FT /tag- c
 FT /product- "Appa2"
 FT /note- "acidic phosphatase"
 FT 241..261
 FT /tag- d
 FT /bound_moiety- "Primer E2"
 FT 1471..1489
 FT /tag- e
 FT /bound_moiety- "Primer K2"
 FT
 PN WO200058481-A2.
 PN
 PD 05-OCT-2000.
 PD
 PF 31-MAR-2000; 2000MO-US08590.
 PF
 PR 31-MAR-1999; 99US-0127032.
 PR
 XX (CORR) CORNELL RES FOUND INC.
 PA
 PI Let X;
 PI
 DR WPI: 2000-619081/59.
 DR P-PDB: AAY94753, AAY94756.
 DR
 PT Novel phosphatase having improved phytase activity useful as animal
 PT feed for improving the accessibility of phosphate to an animal is
 PT produced by treating phosphatase with protease
 PT
 XX
 PS Disclosure; Fig 6; 48pp; English.
 PS
 XX
 CC This invention relates to a phosphatase fragment with improved phytase
 CC activity. The phosphatase fragment is created by treating the phosphatase
 CC with a protease. The invention includes an Escherichia coli nucleotide
 CC sequence which encodes an acidic phosphatase appa2 which has improved
 CC phosphatase activity. The appa gene of Escherichia coli was originally
 CC defined as E. coli periplasmic phosphonhydrolase phosphonhydrolase,
 CC although it was not previously known to have phytase activity. The
 CC phosphatase fragment can be used in animal feed, to improve the
 CC accessibility of phosphate to the animal. The phosphatase gene which
 CC encodes the protein can be used in the production of phytase. The present
 CC sequence represents the appa2 coding sequence.

XX
 SQ Sequence 1489 BP; 362 A; 384 C; 403 G; 340 T; 0 other;
 Alignment Scores:
 Pred. No.: 1.9e-194 Length: 1489
 Score: 2221.00 Matches: 428
 Percent Similarity: 99.07% Conservative: 0
 Best Local Similarity: 99.07% Mismatches: 4
 Query Match: 96.48% Indels: 0
 DB: 21 gaps: 0
 US-09-777-566a-2 (1-440) x AAA28216 (1-1489)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 182 ATGAAAGCGATTCCTTAATCCATTTTATCTTTGATTCCTTAACCCCGCAATCTGCA 241
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArgHisGly 40
 DB 242 TTGCTCAGAGTGAGCCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACCCGCTCAGT 301
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 302 GTGCGTGCCCAACCAAGCCACGCAACTGATGACAGATGTACCCAGACGATGGCCA 361
 QY 61 ThrTriProValLysLeuGlyTyrPleuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 DB 362 ACCGTGGCGGTAAACGTGGTGGCTGACACACCGCGGTGTAGCTTAATCGCCTATCTC 421
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 422 GGAATTATACCAAGCCAGCGCTGTGTGGCCGAGAGATGTGTGGCAAAAAGGCTGGCCG 481
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 482 CAGCTGTGTCAGGTGCGGATTAATGCTGATGTGACAGCGTACCCTAAACAGCGCAA 541
 QY 121 AlaPheAlaIleGlyLeuAlaIleProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 542 GCCCTGCGCCCGCGCGCGCGCTGACTGTGCATTAACCTGTACATCCAGGAGATAGCG 601
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 602 TCCAGTCCGATCCGTTATTTAATCCCTTAATAAAGTGGCTTGGCCAACTGATTAACGCG 661
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
 DB 662 AACGTGACTAGCGGATCTCAGCAGGCGAGAGGTCATTTGCTGACTTACCGGGCAT 721
 QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 722 CGGCAAAAGCGGTTTCCGCAACTGGAACGGGTCTTAATTTTCCCAATTAACCTTGTGC 781
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 782 CTTAACCGTGAAGAAAGCAAGCAAGCTGTTCATTAAACGACGATTAACATGGAACATC 841
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 842 AAGGTGAGCCCGCAACATGTTTCATTAACCGGTGCGGTAAACCTCCGATCAATGATGACG 901
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglArgIleThr 260
 DB 902 GAAATATTTTCTCTGCAACAAAGCAAGCAAGGATCCGAGCGGGGTGGGAAGATCACT 961
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 DB 962 GATTACACCAAGGGAACACCTTCTTAAGTTTGCATTAACCGCAATTTTATTACATCA 1021
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 1022 CGCAGCGCAGAGGTGGCCGCAAGTGGCGCACCCCGCTTATGATTTGATCAATGACGAGCG 1081


```

Db 668 AACGTGACTGACGGCATCTCAGCAGGCGCAGAGGCTCAATTGCTGACTTTACCGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPhenProGlnSerAsnLeuGly 200
Db 728 CGGAAAGCGCGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAAGCTGAGAAACAGACGAAAGCTGTTTCATTAACGACGATTAACCTGGAACCTC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCCCGCCCAATGTTTCATTAACCGGTGCTGTAAGCCCTGCATCAATCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTyrProGlyArgIleThr 260
Db 908 GAAATATTTCTCTCCGCAACAGACAGAGGAAATGCCGAGGCGGAGGAGGAGATCACT 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTCACACACAGTGAACACCTTCTGAAGTTTGATTAACGCGCAATTTATTACTTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCAGCGCCAGAGGTGGCCGACGTGCGCACCCCGTTATTGATTCAGACAGACAGCG 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCCATCCACCCGCAAAACAGCGGTATGTGTGACATTAACCACTTCAGTGTG 1147
Qy 321 PheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATTTCCGAGACAGATTAATCTGCAAAATCTGCGGCGGCGACGAGACTCACTGCG 1207
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1208 ACCGTTCCAGTACGCGGATTAACACCCCGCCAGGTGTGATGATGTTGAACGCTGCG 1267
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGCTAGACGATTAACAGCCAGTGAATTCAGTTCGCTGCTTCACAGACTTTCACAG 1327
Qy 381 GlnMetArgAspLysTrpProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1328 CAGATGCGGTATAAAGCGCGCTATCATTAATACGCGCGCCGAGAGGTGAACCTGACC 1387
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGAGATGTGAAGCGAAATGCCAGGCGATGTGTTGCGTGGCGGTTTAAACGCA 1447
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCGGCATACCGCGCTGCACTTTG 1483

```

RESULT 14
AAC68299
ID AAC68299 standard; DNA; 5421 BP.
XX

AC AAC68299;
XX
XX 20-FEB-2001 (first entry)
XX
DE SV40/APPA plasmid coding sequence.
XX
XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM environmental pollution; pig; ds.
XX
OS Rhesus macaque polyoma virus - chimeric.
XX
OS Escherichia coli - chimeric.
XX
XX NC0200064247-Al.
XX
XX
PD 02-NOV-2000.

```

XX 20-APR-2000; 2000MO-CA00430.
PF
XX 23-APR-1999; 9905-0130508.
PR
XX (UYGU-) UNIV GUELPH.
XX
PI Forsberg CW, Golovan S, Phillips JP;
XX
DR WPI: 2000-687245/67.
DR P-PSDB; AAB36262.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
PS Claim 56; Fig 22; 152pp; English.
CC
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
SO Sequence 5421 BP; 1413 A; 1321 C; 1331 G; 1355 T; 1 other;

Alignment Scores:
Pred. No.: 3,35e-192 Length: 5421
Score: 2205.00 Matches: 429
Percent Similarity: 99.31% Conservative: 429
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 95.79% Indels: 1
Db: 21 Gaps: 0

US-09-777-566a-2 (1-440) x AAC68299 (1-5421)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 40 ATGAAGGCATCTTAATCCATTTTATCTTGTGATCCGTTAACCCGCCAAATGCA 99
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 100 TTGCTCAGATGTAGCCGAGACTGAGCTGAAGTGTGATGTCACGTGCATGCT 159
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 160 GTGGGTGCTCCAAACAGGCGCCAGCACTGATGACAGATGTCACCCAGACGATGGCCA 219
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrLeu 80
Db 220 ACCTGGCGGTAAGACTGGGTGGCTGACACCGCGGNGTGGTGAATCGCTATCTC 279
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 280 GGACATTAACCAAGCCGAGCGCTGTGAGCCGACGATTCGCGGGAAGAGCGTGGCCG 339
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 340 CAGTGTGTCAGTGTGCGATTAATGTCATGATGTCAGACAGCGCTACCCGTAACAGCGAA 399
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 400 GCGTTGCGCGCGGCGGTGGACCTGACTGTGCAATTAACCGTACATACAGGCGATACG 459
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 460 TTCAGTCCCGATCCGTTATTTAATCCTCTAATAAACTGGCGTTGCCAATCGAATAGCG 519
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 520 AACGTGACTGACGGCATCTCAGCAGGCGCAGAGGCTCAATTGCTGACTTTACCGGCAT 579

```

```

OY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGys 200
    |||||||
DB 580 CGCGAAAGCGGCTTCCGCAACTGGAGCGGGTCTTAATTTTCCCAATCAAACTGTGAC 639
OY 201 LeuYsaArgGluYsaGlnAspGluSerGysSerLeuThrGlnAlaLeuProSerGluLeu 220
    |||||||
DB 640 CTTAAACGTGAAGAACAGGACGAAAGCTGTTCAATTAAACGAGGCAATTCATCCGGAACCTC 699
OY 221 LysValSerAlaAspAsnValSerLeuThrGlnValAlaSerLeuAlaSerMetLeuThr 240
    |||||||
DB 700 AAGGAGACCGCCGCAATGTCTCTTAACCGGTGGGTGAAGCTTGCAATGCTGACG 759
OY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglYArgIleThr 260
    |||||||
DB 760 GAGATATTTCTCTGCAACAAGACAGGGAATGCCGAGCGGGGTGGGAAGATCAACC 819
OY 261 AspSerHisGlnTyrPasnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuGln 280
    |||||||
DB 820 GATTCACACGACGAGAACACCTTGCTAGTTGCTTAACGCGCAATTTATTTGCTACAA 879
OY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
    |||||||
DB 880 CGCAGCGCAGAGTTGCCCGCAGCGCCGACCCGTTATTAATGATTCACAGACAGC 939
OY 301 LeuThrProHisProProGlnGlnAlaTyrGlyValIThrLeuProThrSerValLeu 320
    |||||||
DB 940 TTGACGCCCCACCA-CCGCAAAACAGCGCATGCTGATGACATTCACCTTCAGTCTG 998
OY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnThr 340
    |||||||
DB 999 TTTATCGCCGACACGATCTAATCTGCAATCTGCGCGGGGCGACTGGAGCTCAACG 1058
OY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluArgTyr 360
    |||||||
DB 1059 ACGCTCCCGGTGACCGCGATTAACAGCGCGCAGGTGTGAACGTGTTGAACGCTGG 1118
OY 361 ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln 380
    |||||||
DB 1119 CGTGGCGTAAACGATTAACAGCGAGTGCAGTTGCTGCTTCAGACTTTACAG 1178
OY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
    |||||||
DB 1179 CAGATGCTGATTAACAGCGCCCTGTCAATTAATACGCCCGCGAGAGTGAACTGACC 1238
OY 401 LeuAlaGlyGysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
    |||||||
DB 1239 CTGGCAGAGTGTGAAGACGAAATGCCAGGAGCATGTGCTGTCGAGGTTTTCGCAA 1298
OY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
    |||||||
DB 1299 ATCGTGAATGAACGACGATACCGCGTTCGAGTTTG 1334

RESULT 15
AAD06832
ID AAD06832 standard; DNA; 1486 BP.
XX
AC AAD06832;
XX
DT 06-AUG-2001 (first entry)
XX
DE E. coli acid phosphatase/phytase (appa) mutant DNA.
XX
KW Acid phosphatase-phytase; appa; enzymatic activity; food additive;
KW animal feed; monooester phosphate; cereal food; human food; mutant; ds.
XX
OS Escherichia coli.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 16..108
FT CDS
FT /tag= a
FT /product= "peptide encoded by shorter ORF of mutant appa

```

```

FT FT CDS DNA"
FT FT 188..1486
FT FT /tag= b
FT FT /product= "E. coli appa mutant"
FT FT /transl_except= (pos:512..514, aa:Ala)
FT FT /transl_except= (pos:629..631, aa:Asn)
FT FT 188..247
FT FT /tag= c
FT FT 248..1483
FT FT /tag= d
FT FT /product= "Mature mutant appa protein"
FT FT replace (785..786, 76)
FT FT /tag= e
FT FT mutation replace (806, G)
FT FT /tag= f
FT FT mutation replace (808, C)
FT FT /tag= g
FT FT mutation replace (818..820, TCA)
FT FT /tag= h
FT FT
FT FT WO200136607-A1.
FT FT
FT FT 25-MAY-2001.
FT FT
FT FT 17-NOV-2000; 2000WO-US31622.
FT FT
FT FT 18-NOV-1999; 99US-0166179.
FT FT
FT FT (CORR ) CORNELL RES FOUND INC.
FT FT
FT FT Let X:
FT FT
FT FT WPI: 2001-367572/38.
FT FT
FT FT P-PSDB; AAE02632, AAE02633.
FT FT
FT FT Mutated acid phosphatase/phytase from Escherichia coli has improved
FT FT enzymatic activity compared to the wild type and is useful as a food
FT FT additive, particularly for animal feeds
FT FT
FT FT Claim 11; Page 11-12; 56pp; English.
FT FT
FT FT
FT FT The present sequence is a DNA encoding mutant acid phosphatase/phytase
FT FT (appa) from Escherichia coli. The wild type appa is mutated to enhance
FT FT its enzymatic activity. Phytases, a specific group of monooester
FT FT phosphates, are required to initiate the release of phosphate from
FT FT phytate, the major storage of phosphate in cereal foods or feeds. The
FT FT mutant acid phosphatase/phytase is added to animal feed as a food
FT FT additive to improve uptake of phosphate and zinc from the diet. The
FT FT phytase can also be added to human food. The mutant phytase is more heat
FT FT stable than phytase isolated from Aspergillus niger and safer for use
FT FT in human food manufacture.
FT FT
FT FT Sequence 1486 BP; 371 A; 379 C; 401 G; 335 T; 0 other;
FT FT
FT FT Alignment Scores:
FT FT Pred. No.: 3,05e-192 Length: 1486
FT FT Score: 2197.00 Matches: 424
FT FT Percent Similarity: 98.61% Conservative: 2
FT FT Best Local Similarity: 98.15% Mismatches: 6
FT FT Query Match: 95.44% Indels: 0
FT FT DB: 22 Gaps: 0
FT FT
FT FT US-09-777-566a-2 (1-440) x AAD06832 (1-1486)
FT FT
OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
    |||||||
DB 188 ATGAAGCGATCTTAATCCATTTTATCTCTTGATTCGTTAACCAGCGCAATCTCA 247
OY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
    |||||||
DB 248 TTCCGCTCAGAGTGAGCCGAGACTGAAGCTGAAGTGTGATGATGACGCGTCATGCT 307
OY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro 60
    |||||||

```

|||||
Db GTGGGTGCCCCAACCAAGGCCACCACTGATGAGATGTCACCCAGACGATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgLysGlyLeuIleAlaTyrLeu 80
Db ACCGTGGCCGGTAAACATGGGTGGTGGTACACACCGCGGTGGTGAATGAGCTTATCTC 427
QY 81 GlyHisTyrGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
Db 428 GGACATTACCAACCCACCGCTGTGTGGCCACGATGCTGGGCAAAAAGGCTGCCG 487
QY 101 GlnSerGlyValAlaIleIleAlaAspValAspGluArgTyrGlyGlu 120
Db 488 CAGCGTGTGTCAGTCGCGATATTGTGATGTCGACGAGCTACCCGTAACAGCGCA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCGTTCGCGCGGGGTGGCCACTGCTGCAATACCGTACATACCCAGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuTyrThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TTCAGTCCCGATCCGTTATTATTCTTAAACCTGGCGTTTGCCACATGATTAACGCG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTAGTACGCGGATCTCAGCAGGAGGAGGATCAATTGCTGACTTACCGGCGCAT 727
QY 181 ArgGlnThrAlaPheArgGlyLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGGCTTCGGAACCTGAAACGGGCTTAAATTTCCGCAATCAAACTTGAC 787
QY 201 LeuTyrArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 788 CTTAAACGTGGAACAGAACTGTAACCTTAACGACGATTAACGAACTC 847
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTTTCATTAACGCGTACGCTGCAATGCTGACG 907
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Db 908 GAAATATTTCTCTGCAACAGACAGGAAATGCGGAGCGGGGTGGGGAAGATCACT 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTCAACACAGTGGAACTTCTGTAAGTTGCTAATACGCGCAATTTTATTACTACAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCAGCGCAGAGTGGCCGAGTCGCGCCACCCCGTTATGATTGATCAAGACAGCG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGGCCCATCCACCGCAAAACAGCGGTATGTGTGACATTACCCACTTCAGTCTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGluLeuAsnTrp 340
Db 1148 TTTATTGCGGACACGATACTAATCTGCAATCTCGCGGCGCATGAGCTCAACTGG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluArgTrp 360
Db 1208 ACGCTTCAGGTACGCGGATACACGCGCCAGGTGTGTAACGTGTTGAACGCTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGGTAAACGATACAGCAAGTGTGATTCAGGTTTCGCTTCACACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1328 CAGATGCGTGAATAAAGCCGCTATCATTAATACGCGCGCGAGAGAGTGAACCTGACC 1387
QY 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
|||||

Db 1388 CTGGCAGGATGTGAAGACGAATAATGCGCAGGCAATGTGTCGTGGCCGTTTACGCA 1447
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTAATGAAGCGCGCATACCGCGCTGCAGTTTG 1483

Search completed: June 12, 2003, 13:27:26
Job time : 408 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 12:11:14 ; Search time 1929 Seconds

(without alignments)
3694.149 Million cell updates/sec

Title: US-09-777-566A-2

Perfect score: 2302

Sequence: 1 MKAILPLSLILPLTPQSA.....IVNEARIPACSLRHHHHH 440

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame_plus_p2n.model -DEV=xlh
-o/cgn2.1/USPTO_Spool/US09777366/runat_05062003_154808_6562/app_query.fasta.1.583
-DB=EST -QFMT=fastap -SUFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-OCALIGN=200 -THR.SCORE=Pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Plc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09777366 -GCN1.1 1906 -etunat_05062003_154808_6582 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-NGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
EST:*
1: em_estba:*
2: em_estbun:*
3: em_estlin:*
4: em_estimu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1044	45.4	616	13	BJ074127 BJ074127
C 2	643	27.9	383	10	AM036132 EST274508
C 3	643	27.9	383	10	AM036134 EST274510
C 4	641	27.8	393	10	BE520240 BE520240
C 5	511.5	22.2	354	9	AA545747 HMFSP184 -
C 6	392	17.0	320	10	BE346403 BE346403
C 7	355	15.4	208	13	BM409016 BM409016
C 8	355	15.4	208	13	BM412806 BM412806
C 9	339	14.7	546	13	BJ040796 BJ040796
C 10	287	12.5	657	12	BG457132 NF100C05P
C 11	278	12.1	179	10	BE461872 EST413291
C 12	269	11.7	508	13	BJ030203 BJ030203
C 13	228.5	9.9	480	9	AL586973 AL586973
C 14	225	9.8	254	10	BE520241 BE520241
C 15	172.5	7.5	696	9	AU004475 AU004475
C 16	144.5	6.3	645	13	BJ037937 BJ037937
C 17	125	5.4	274	14	BO153360 NF035C07I
C 18	119.5	5.2	557	13	BJ029368 BJ029368
C 19	118	5.1	1075	14	BM922358 AGENCOURT
C 20	116.5	5.0	720	13	BM628471 170006874
C 21	114.5	5.0	702	12	BF037418 601460789
C 22	114	5.0	1010	13	BM465775 AGENCOURT
C 23	112.5	4.9	1060	14	BM922665 AGENCOURT
C 24	112	4.9	1797	11	BC013495 Mus muscu
C 25	109.5	4.8	1080	11	BQ278048 AGENCOURT
C 26	109	4.7	2396	11	BC022504 Homo sapi
C 27	108.5	4.7	665	13	BI764510 603050802
C 28	108	4.7	796	10	AV400670 AV400670
C 29	107.5	4.7	657	13	BI265620 NF096B07I
C 30	107	4.6	686	12	BG457966 NF036A10P
C 31	107	4.6	842	12	BG026657 AGENCOURT
C 32	107	4.6	1114	9	BM920972 AL545650
C 33	105.5	4.6	915	9	AL545650
C 34	105	4.6	567	10	AM666304 AGENCOURT
C 35	104.5	4.5	1129	14	BM926193 AGENCOURT
C 36	104.5	4.5	1212	13	BM547032 AGENCOURT
C 37	104	4.5	565	13	BM084955 sa350h07
C 38	103.5	4.5	775	13	BI819484 603034283
C 39	103	4.5	548	12	BF519356 EST45618
C 40	103	4.5	924	12	BE889074 601513513
C 41	102.5	4.5	1907	12	BF166252 60177014
C 42	102	4.4	621	10	AM175388 f135e03.y
C 43	102	4.4	1042	14	BM903520 AGENCOURT
C 44	101.5	4.4	1161	14	BM926491 AGENCOURT
C 45	101.5	4.4	1197	13	BI489085 603021120

ALIGNMENTS

```
RESULT 1
BJ074127/c 616 bp mRNA linear EST 11-DEC-2001
LOCUS BJ074127 N1BB Mochii normalized xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone X1090106 5', mRNA sequence.
ACCESSION BJ074127
VERSION BJ074127.1 GI:17504316
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 616)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara
```

TITLE
JOURNAL
COMMENT

Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadashi Shinn-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshinn@genes.nig.ac.jp

FEATURES
source

Location/Qualifiers
1.616
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="X1090106"
/clone_lib="NIBB Mochi normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Mieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 128 a 167 c 167 g 154 t
ORIGIN

Alignment Scores:

Pred. No.: 7.59e-100 Length: 616
Score: 1044.00 Matches: 204
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 45.35% Indels: 0
DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BJ074127 (1-616)

QY 89 ValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIle 108
DB 615 GTAGCCGACGAGATTGCTGGCGAATAAGCGCTCCCGACACTGTCAGTCGCGGATATT 556
QY 109 AlaAspValAspGlnArgThrArgLysThrGlyGlnAlaAlaGlyLeuAlaPro 128
DB 555 GCTGATGTCACAGCGGTACCGCTAAACAGCGGAAAGCTTCCCGCGGGCTGGACCT 496
QY 129 AspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeuAsn 148
DB 495 GACTGTGCATATACCTACATACCCAGCGAGATACCTCCAGTCCGCTATTATTAT 436
QY 149 ProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSer 168
DB 435 CCTCTAAAACCTGGCTTCCCAACTGATTAACGCAACCTGACTGACGCGATCTCAGC 376
QY 169 ArgAlaGlyGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGlnLeu 188
DB 375 AGGCGAGAGGAGCAATTGCTGACTTACGGGCAATGGGAAAGCGGCTTCCGGAATCG 316
QY 189 GlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGlnLysGlnAspGln 208
DB 315 GAACGGGTCTTAATTTTCCGCAATCAACTTGTCTTAAAGTGAGAAACAGAGCA 256
QY 209 SerCysSerLeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspAsnValSer 228
DB 255 AGCTGTTCATTAACGACGACATTCATCGAAGCTCAAGGTGAGCCCGCAATGTCTCA 196
QY 229 LeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuGlnGlnAla 248
DB 195 TTAAACGGGTGGTAAAGCTTCGATCAATGCTGACGAGATATTCTTCCGCAACAGCA 136
QY 249 GlnGlyMetProGlnProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeu 268
DB 135 CTGGGAATGCGGAGCGGGGTGGGAGAGATCAACGATTACACCAAGTGGAACACCTTG 76

QY 269 LeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGlnArgThrProGlnValAlaArgSer 288
DB 75 CTAAAGTTGATTAACGCGCAATTTATTCTCAACAGCACGCGAGGTTGCCCGACG 16
QY 289 ArgAlaThrProLeu 293
DB 15 CGCGCACCCCGTTA 1

RESULT 2
AM036132 383 bp mRNA linear EST 18-MAY-2001
LOCUS
DEFINITION
EST274508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
cDNA1E23 similar to periplasmic phosphonhydrolase precursor, putative, mRNA sequence.

ACCESSION
AM036132
VERSION
AM036132.1 GI:5894811
KEYWORDS
EST.
SOURCE
ORGANISM
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 383)
Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Roming,C.M., Craven,M.B., Fujii,C.Y., Bowman
C.L., Nierman,M., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
S.D. and Giovannoni,J.
Generation of ESTs from tomato seed tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
3 prime sequence.

REFERENCE
AUTHORS
1.383
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cDNA1E23"
/clone_lib="tomato seed, TAMU"
/tissue_type="seeds"
/dev_stage="quiescent seed"
/lab_host="X11-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cDNA - Tomato seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."

FEATURES
source
Location/Qualifiers
1.383
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cDNA1E23"
/clone_lib="tomato seed, TAMU"
/tissue_type="seeds"
/dev_stage="quiescent seed"
/lab_host="X11-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cDNA - Tomato seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."

BASE COUNT 102 a 101 c 97 g 83 t
ORIGIN

Alignment Scores:
Pred. No.: 8.28e-58 Length: 383
Score: 643.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.93% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x AM036132 (1-383)

QY 117 LysThrGlyGlnAlaPheAlaGlyLeuAlaProAspCysAlaIleThrValHisThr 136
DB 2 AAAACAGCGCAACCTTCCGCGCGGCTGGACCTGACCTGACATTAACCGTACATACC 61
QY 137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln 156
DB 62 CAGGAGATACGTCACGATCCGATTCGATTAATTCCTTAATAACTGGCGTTGGCAA 121
QY 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAsp 176

Db 122 CTGATTAACGCCAGCTGATGACCGCATCTCAGACGAGGAGAGGATGCAATTCGTGAC 181
 QY 177 PheThrgLysIArgGlnThraLpheArgGluLeuGluArgValLeuAsnHepGln 196
 Db 182 TTTACCGGCGATCGCAAAAGCGGCTTCCGCAACGTGAAAGCGGCTTAAATTTCCGCAA 241
 QY 197 SerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerLeuThrgLysAlaLeu 216
 Db 242 TCAAACTGTGCTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGACATTA 301
 QY 217 ProSerGluLeuLysValSerAlaAspAsnValSerLeuThrgLysAlaValSerLeuAla 236
 Db 302 CCATCGGAACCTCAAGCTGAGCGCCGACATGTCATTAACCGGTGGTAAGCTTCGCA 361
 QY 237 SerMetLeuThrgLysIlePhe 243
 Db 362 TCAATGCTGACGAGATATTT 382
 RESULT 3
 AM036134/c
 LOCUS 383 bp mRNA linear EST 18-MAY-2001
 DEFINITION ES2724510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
 CLE1E23 similar to periplasmic phosphonhydrolase precursor, putative, mRNA sequence.
 ACCESSION AM036134
 VERSION AM036134.1 GI:5894813
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 REFERENCE 1 (bases 1 to 383)
 Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
 ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
 ,S.D. and Giovannoni,J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 JOURNAL Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 COMMENT 5 prime sequence.
 FEATURES
 source
 1..383
 /organism="Lycopersicon esculentum"
 /cultivar="TAM496"
 /db_xref="taxon:4081"
 /clone="CLE1E23"
 /clone_lib="tomato seed, TAMU"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
 XhoI; cDNA - Tomato Seed EST Library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."
 BASE COUNT 83 a 97 c 101 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.28e-58 Length: 383
 Score: 643.00 Matches: 127
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.93% Indels: 0
 Gaps: 0
 US-09-777-566a-2 (1-440) x AM036134 (1-383)

QY 117 LysThrgLysIArgGlnThraLpheArgGluLeuGluArgValLeuAsnHepGln 136
 Db 382 AAAACAGCGGAAGCCTTCCCGCGGCTGCGACCTGACTGTGCATTAACCGTACATACC 323
 QY 137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrgLysAlaCysGln 156
 Db 322 CAGGCGAGATACGTCGAGTCCCGATCCGTTAATTAATTCCTTAAAAACGTGGCTTGGCCA 263
 QY 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAsp 176
 Db 262 CTGATTAACGCAACCTGACTGACCGGATCTCAGCAGGAGAGAGGTCATTCGTGAC 203
 QY 177 PheThrgLysIArgGlnThraLpheArgGluLeuGluArgValLeuAsnHepGln 196
 Db 202 TTTACCGGCGATCGCAAAAGCGGCTTCCGCAACCTGGAACGGGTGTTAATTTCCGCAA 143
 QY 197 SerAsnLeuCysLeuLysArgGluLysGlnAspGluSerCysSerLeuThrgLysAlaLeu 216
 Db 142 TCAAACTGTGCTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGACATTA 83
 QY 217 ProSerGluLeuLysValSerAlaAspAsnValSerLeuThrgLysAlaValSerLeuAla 236
 Db 82 CCATCGGAACCTCAAGCTGAGCGCCGACATGTCATTAACCGGTGGTAAGCTTCGCA 23
 QY 237 SerMetLeuThrgLysIlePhe 243
 Db 22 TCAATGCTGACGAGATATTT 2
 RESULT 4
 BE520240/c
 LOCUS 393 bp mRNA linear EST 19-MAR-2001
 DEFINITION M1B125NM Arabidopsis developing seed Arabidopsis thaliana cDNA
 clone M1B12 5', mRNA sequence.
 ACCESSION BE520240
 VERSION BE520240.1 GI:9778242
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 393)
 White,J.A., Todd,J., Newman,T., Focks,N., Girke,T., Martinez de
 Alarduya,O., Jaworski,J.G., Ohlroge,J. and Benke,C.
 A new set of Arabidopsis expressed sequence tags from developing
 seeds. The metabolic pathway from carbohydrates to seed oil
 Plant Physiol. 124 (4), 1582-1594 (2000)
 JOURNAL Contact: Benning, C
 MEDLINE Dept. of Biochemistry & Molecular Biology
 COMMENT Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 , USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TEL: 6142929371.
 FEATURES
 source
 1..393
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="M1B12"
 /clone_lib="Arabidopsis developing seed"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 86 a 100 c 103 g 102 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 1,4e-57 Length: 393
 Score: 641.00 Matches: 127
 Percent Similarity: 97.71% Conservative: 1
 Best Local Similarity: 96.95% Mismatches: 3
 Query Match: 27.85% Indels: 0
 DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x BE520240 (1-393)

QY 107 ILeIleAlaAspValaspGluArgThrArgLysThrGlyLysAlaPheAlaGlyLeu 126
 Db 393 ATTATTGCTGATGTCGACGACGCTACCCGTAANACGACCAACTTCCGCCGGGNTG 334
 QY 127 AlaProAspCysAlaIleThrValHisThrGlnAlaAspHisSerProAspProLeu 146
 Db 333 GCACCTGACTGTGCANATACCGTACATACCCAGCAGATACGTCAGTCCGATCCGTTA 274
 QY 147 PheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIle 166
 Db 273 TTTAAATCCTCTAAAACTGCGCTTGCCACCTGATTAACGCGAAGTGAACGCGATC 214
 QY 167 LeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArg 186
 Db 213 CTCAGCAGGCGACGAGGCTCAATTGCTGACTTACCGGCGCATCGCGAAGCGGCTTCCG 154
 QY 187 GlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGln 206
 Db 153 GAACCTGACGCGGTGCTTAATTTTCCGCAATCAAACTTGCTTAACTGAGAAACAG 94
 QY 207 AspGluSerCysSerLeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspAsn 226
 Db 93 GACGAAGCGTTCATTAAACGACGACGATTCACATCGGAACCTCAAGGTGAGCGCGACAT 34
 QY 227 ValSerLeuThrGlyAlaValSerLeuAlaSer 237
 Db 33 GTCTCATTAAGCGCTGCGGTAAACCTTCGCATCA 1

RESULT 5 AA545747 354 bp mRNA linear EST 12-MAY-1999
 AA545747/c LOCUS
 DEFINITION HBMS1B4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens cDNA
 clone HBMS1B4 5', mRNA sequence.
 ACCESSION AA545747
 VERSION AA545747.1 GI:2307026
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Jia,L., Robey,P., Young,M., Schuler,G., Powell,J., Yang,L., Lennon
 G., Hillier,L., Allen,M., Bowles,L., Geisel,S., Kucaba,T., Marra
 M., Martin,J., Steptoe,M., Tan,F., Theising,B., Bowers,Y., Wylie
 T., Waterston,R., Wilson,R. and Francomano,C.
 TITLE WashU-MGB/NHGRI EST Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Libin Jia
 National Genetics Branch
 Medical Genetics Research Institute
 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
 Tel: 301-402-4877
 Fax: 301-496-7157
 Email: libin@helix.nih.gov
 Seq primer: M13 Reverse.
 Location/Qualifiers

FEATURES
 source 1..354
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HBMS1B4"
 /clone_1lb="Human Bone Marrow Stromal Fibroblast"

/sex="Male and Female"
 /tissue_type="bone marrow"
 /cell_type="stromal fibroblast"
 /dev_stage="mixed"
 /lab_host="XLI-blue"
 /note="Vector: pBluescript; Site_1: EcoRI; Site_2: XhoI"

Alignment Scores:

Pred. No.: 5.71e-44 Length: 354
 Score: 511.50 Matches: 112
 Percent Similarity: 94.96% Conservative: 1
 Best Local Similarity: 94.12% Mismatches: 5
 Query Match: 22.22% Indels: 4
 DB: 9 Gaps: 0

US-09-777-566a-2 (1-440) x AA545747 (1-354)

QY 166 ILeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPhe 185
 Db 352 GTCTTCAGACGAGGACGAGGTCA-ATTGCTGA-TTTACGGGCGATCGGCAAGCGGCTT 295
 QY 185 eArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuLysArgGln 205
 Db 294 TCCGGAACCTGGAACGGGTGTT-AATTTCCGCAATCAAACTTGCTTAAAGTGAGAA 236
 QY 205 sGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAs 225
 Db 235 ACAGGACGAAGCGTTCATTAAACGACGACATTCATCCATCGGAACCTCAAGGTGAGCGG 176
 QY 225 PAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnLeuPheLeu 245
 Db 175 CAATGTTCATTAACCGGTGCGGTAAAGCCGTCATCAATCTGACGAGATATTTCCT 116
 QY 245 uGlnGlnAlaGlnGlyMetProGlnProGlyTTPGlyArgIleThrAspSerHisGlnTr 265
 Db 115 GCAACAAGCAGACGGAATGCCGAGCGGGGTGGGAGAGATACCGGATTCACACAGTG 56
 QY 265 PAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuGlnArgThrPro 283
 Db 55 GAACACCTTGCTAAGTTTGATATACCGGCAATTTATTCTCTACACGCGCGCA 1

RESULT 6 BE436403 320 bp mRNA linear EST 18-MAY-2001
 BE436403/c LOCUS
 DEFINITION EST407481 tomato breaker fruit, T1R Lycopersicon esculentum cDNA
 clone cLEG32C20, mRNA sequence.
 ACCESSION BE436403
 VERSION BE436403.1 GI:9434246
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronsing,C.M.,
 Niernan,M., Fraser,C.N., Martin,G.B., Giovannoni,J.J. and Tanksley
 S.D.
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage
 JOURNAL Unpublished (2000)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 Location/Qualifiers

FEATURES
 source 1..320
 /organism="Lycopersicon esculentum"

100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics Institute
 Seq primer: T3.

FEATURES

source

1. 208 location/Qualifiers

/organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CDE61013"
 /clone_1lb="tomato breaker fruit"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCunadapt; Site_1: EcoRI; Site_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 43 a 56 c 62 g 47 t

ORIGIN

Alignment Scores:
 Pred. No.: 8.37e-28 Length: 208
 Score: 355.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.42% Indels: 0
 DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BM412806 (1-208)

QY 82 HistyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIstysGlyCysProGln 101
 |||||
 Db 206 CATTACCAACGCCAGCTCTGCTAGCCGACGATGCTGCGAAGGAGGCTGCCAG 147
 QY 102 SerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIstysGlyAla 121
 |||||
 Db 146 TCTGTCAGGTCGCGATTATTGCTGATGTCGACGAGGTACCGGTAAACAGGAGGCC 87
 QY 122 PhAlAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSer 141
 |||||
 Db 86 TTGCGCGCGGCGTGCACCTGACTGCAATTAACGTAACCCAGGAGATACGCTCC 27
 QY 142 SerProAspProLeuPheAsnPro 149
 |||||
 Db 26 AGTCCGATCCGTTATTATTAATCT 3

RESULT 9
 BJO40796 546 bp mRNA linear EST 06-DEC-2001
 LOCUS BJO40796 NIBB Mochii normalized Xenopus neurula library xenopus
 DEFINITION laevis cDNA clone XLO47108 5', mRNA sequence.

ACCESSION BJO40796
 VERSION BJO40796.1 GI:17388187
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 546)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsuhin@genes.nig.ac.jp
 location/Qualifiers

FEATURES

source

1. 546 location/Qualifiers

/organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XLO47108"
 /clone_1lb="NIBB Mochii normalized Xenopus neurula library"
 /tissue_type="whole embryo"
 /dev_stage="stage 15"
 /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAs were oligo-dT primed and directionally cloned. Staging according to Mieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 116 a 145 c 142 g 138 t 5 others

ORIGIN

Alignment Scores:
 Pred. No.: 1.92e-25 Length: 546
 Score: 339.00 Matches: 66
 Percent Similarity: 98.51% Conservative: 0
 Best Local Similarity: 98.51% Mismatches: 1
 Query Match: 14.73% Indels: 0
 DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BJO40796 (1-546)

QY 138 AlaAspThrSerSerProAspProLeuPheAsnProLeuYsthrGlyValCysGlnLeu 157
 |||||
 Db 345 GCGGTGACGTCACAGTCCGATCCGATTTAAATCTTAACCTGAAACCTGGCTTGCAACTG 404
 QY 158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPhe 177
 |||||
 Db 405 GATACCGCAACGATGATGACCATCTCAGACGAGGAGGTAAATTCGACTTT 464
 QY 178 ThrGlyHisArgGlnThrAlaPheArgGlnLeuGlnValArgValLeuAsnPheProGlnSer 197
 |||||
 Db 465 ACCGGGATCGGCAACAGGCGTTTCCGAACTGGAAAGCGGTATTTTCCGAAATCA 524
 QY 198 AsnLeuCysLeuYsthrGlu 204
 |||||
 Db 525 AACTGTGCTTAACGAG 545

RESULT 10
 BG457132 657 bp mRNA linear EST 19-MAR-2001
 LOCUS BG457132
 DEFINITION NF100C05P1F1036 Phosphate starved leaf Medicago truncatula cDNA
 ACCESSION BG457132
 VERSION BG457132.1 GI:13380457
 KEYWORDS EST.

SOURCE barcel medic.
 ORGANISM Medicago truncatula

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 657)
 Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 Medicago truncatula phosphate-starved leaf library
 Unpublished (2000)

TITLE Unpublished (2000)
 JOURNAL Contact: Harrison MJ
 COMMENT Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7325
 Fax: 580 221 7380
 Email: mjharrison@noble.org

/cultivar="TA496"

According to Nieuwkoop and Faber. Library is subtracted

and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute).
BASE COUNT 108 a 137 c 137 g 126 t
ORIGIN

Alignment Scores:

Pred. No.: 4,26e-18 Length: 508
Score: 269.00 Matches: 53
Percent Similarity: 98.15% Conservative: 0
Best Local Similarity: 98.15% Mismatches: 1
Query Match: 11.69% Indels: 0
DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BU030203 (1-508)

QY 138 AlaAspThrSerSerProAspProLeuPheAsnProLeuThrGlyValCysGlnLeu 157
Db 345 GCGGTACGTCACGTCGCCGATCGCTATTATTCCTCTAAAGTGGCGTTGCCAACTG 404
QY 158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPhe 177
Db 405 GATTAACGCGAACCTGACTGACGCGATCCACAGGAGGAGGTCATTCGACTTT 464
QY 178 ThrGlyHisArgIleThrAlaPheArgGluLeuGluArgVal 191
Db 465 ACCGGCATCGCAACGCGCTTTCGCAACTGGAACGGGTG 506

RESULT 13
LOCUS AL586973 480 bp mRNA linear EST 02-MAR-2001

DEFINITION AL586973 Stratagene Chick Embryo Lambda cDNA Library (* 93/405)
Gallus gallus cDNA clone ROS050E02, mRNA sequence.

ACCESSION AL586973
VERSION AL586973.1 GI:13192007
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
1 (bases 1 to 480)

AUTHORS Murray, F.
TITLE Stratagene Chick Embryo Lambda cDNA Library
JOURNAL Unpublished (2001)
COMMENT Contact: Frazer Murray
Dept. Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UK
Tel: +44 (0)131 527 4200
Fax: +44 (0)131 440 0434
Email: frazer.murray@bbsrc.ac.uk
Seq primer: T3.

FEATURES

source 1. 480 Location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="ROS050E02"
/clone_lib="Stratagene Chick Embryo Lambda cDNA library (* 93/405)"
/tissue_type="Embryo"
/dev_stage="5 days Old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
; Cloned unidirectionally. Primer: Oligo dt. Uni-ZAP XR
vector. Average insert size: 1.5kb; 5' adaptor sequence:
5' GAATTCGCGACGAG 3'; 3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3'"
BASE COUNT 111 a 129 c 114 g 125 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 7.39e-14 Length: 480
Score: 228.50 Matches: 58

Percent Similarity: 50.29% Conservative: 28
Best Local Similarity: 33.92% Mismatches: 68
Query Match: 9.93% Indels: 17
DB: 9 Gaps: 4

US-09-777-566a-2 (1-440) x AL586973 (1-480)

QY 235 LeuAlaSerMetLeuThrGluIlePheLeuGlnGlnAlaGlnGlyMetPro----- 252
Db 480 GTCGGCACTCGCTGGTATGCGTTTACTTTCGAATATTCAGAAAGTTTCCGATGAT 421
QY 253 GluProGlyThrPheGlyGlyIleThrAspSerHisGlnThrAsnThrLeuLeuSerLeuHis 272
Db 420 CAGGTGGCTGGGAGAGAAATCAATCTGACACGACGATGAGGTGTTGTCGAAGCTTAAA 361
QY 273 AsnAlaGlnPheThrLeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrPro 292
Db 360 AACGGCTACCAAGACGACCTGTTTACTCTACCGGAAGTGGCGCAATGTGGCAACCG 301
QY 293 LeuLeuAspLeuIleMetAlaAlaLeuThrProHisProGluGlnGlnAlaThrGly 312
Db 300 CTGGTCAGTATATTCGACAAAGCTCTG----- 274
QY 313 ValThrLeuProThrSer-----ValLeuPheIleAlaGlyHisAspThrAsnLeu 329
Db 273 GTCACCGATCGACACGACGCGAAATTAACAGTGTGGTGGCGACGCTCCAACTT 214
QY 330 AlaAsnLeuGlyGlyAlaLeuGlnLeuAsn---TriThrLeuProGluGlnProAspAsn 348
Db 213 GCCTCTCTGTAAACGGCGTGGATTTCAACCGTATGATGATGATGATGATGATGATGATGAT 154
QY 349 ThrProGluGlyGlyLeuValPheGluArgThrArgThrArgLeuSerAspAsnSerGln 368
Db 153 ACCCGCATGCGCGCAAAATCGTTTCCACGCTGGCATGACAGCAAAAGCAATCGCAT 94
QY 369 TriPheGlnValSerLeuValPheGlnThrLeuGlnGlnMetArgAspThrProLeu 388
Db 93 TTGATGAATTAATTAATGATGATACAGAGTGGCAACGATTAATGATGATGATGATGATGAT 34
QY 389 SerLeuAsnThrProProGlyGluValLeuLeu 399
Db 33 ACCCTGACGACCTCTGCAAGCGTGTGACCTG 1

RESULT 14
LOCUS BE520241 254 bp mRNA linear EST 19-MAR-2001
DEFINITION M1B12XTM Arabidopsis developing seed Arabidopsis thaliana cDNA
clone M1B12 3', mRNA sequence.
ACCESSION BE520241
VERSION BE520241.1 GI:9778243
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops;
1 (bases 1 to 254)
AUTHORS White, J.A., Todd, J., Newman, T., Focks, N., Gilke, T., Martinez de
Ibarra, O., Jaworski, J.G., Ohlrogge, J. and Bensing, C.
A new set of Arabidopsis expressed sequence tags from developing
seeds. The metabolic pathway from carbohydrates to seed oil
Plant Physiol. 124 (4), 1582-1594 (2000)
20567808
CONTACT: Bensing, C
Dept. of Biochemistry & Molecular Biology
Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824
, USA

Tel: 517 355 1609
Fax: 517 353 9334
Email: bensing@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 303 Botany &

Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
6142920603 TEL: 6142929371.

FEATURES

Location/Qualifiers
1..254

ORGANISM "Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="M1B12"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E.coli"
/note="Organ: Developing seed; Vector: pbluescript SK-;
Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 70 a 57 c 63 g 64 t
ORIGIN

Alignment Scores:

Pred. No.: 6,26e-14 Length: 254
Score: 225.00 Matches: 47
Percent Similarity: 92.31% Conservative: 1
Best Local Similarity: 90.38% Mismatches: 4
Query Match: 9,778 Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x BE520241 (1-254)

OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
DB 96 ATGAAAGCATCTTATCCCATTTTATCTCTGATTCCTTAACCCCGCAATCTGCA 155
OY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
DB 156 TTCGCTCAGAGGAGCCGAGCTGAAGCTGGAAGTGTGCTATGTCTCAGTGTCTATG 215
OY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGln 52
DB 216 GTGCTGTCTCCACCAAGCAGCACTGATGATGAG 251

RESULT 15

AU004475

LOCUS AU004475 Bombyx mori p50(Daizo) Bombyx mori cDNA clone ws20511,
DEFINITION mRNA sequence.
ACCESSION AU004475
VERSION AU004475.1 GI:4161846
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 696)
AUTHORS Mita,K., Morimoto,M., Shimada,T., Okano,K. and Maeda,S.
TITLE Establishment of cDNA database of Bombyx mori
JOURNAL Unpublished (1999)
COMMENT Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmila@irs.go.jp
PROJECT "CREST project by JST".

FEATURES

Location/Qualifiers
1..696

ORGANISM "Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="ws20511"
/clone_lib="Bombyx mori p50(Daizo)"
BASE COUNT 237 a 132 c 136 g 191 t
ORIGIN

Alignment Scores:

Pred. No.: 1.11e-07 Length: 696
Score: 172.50 Matches: 49
Percent Similarity: 44.69% Conservative: 31
Best Local Similarity: 27.37% Mismatches: 78
Query Match: 7,49% Indels: 21
DB: 9 Gaps: 5

US-09-777-566a-2 (1-440) x AU004475 (1-696)

OY 28 LeuLysLeuGlnSerValIleValSerArgHisGlyValArgAlaProThrLysAla 47
DB 130 CTACGCTTGACCAAGTATTATTTTACTCGTCAACACTTAAGAGTTCCA-----CTA 183
OY 48 ThrGlnLeuMetGlnAspValThrProAspAlaThrProThrTyrProValLysLeuGly 67
DB 184 TCATAACATATGATGATGATACACGACAAATAATTTCCGAATGGTCCCAAGCCGGCC 243
OY 68 TrpLeuThrProArgGly-----GlyGluLeuIleAlaTyrLeuGlnHisTyr 83
DB 244 ATGCTGACGCAAAAGGACCTTATTTGAAAGTCATATGAGGAATATCTAAGCAATGG 303
OY 84 GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProGlnSerGly 103
DB 304 ATGAGAGAGAAATCAACTTTTACCCGCAAGC-----TGTCGCAACAAAGA 348
OY 104 GlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGluAlaPheAla 123
DB 349 ACCGTTTGTATTTATGCGATATTAAGACACACACCATACCTGCTGCGAAGCATTTGTC 408
OY 124 AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerPro 143
DB 409 GATGCCGATTCCTCCGACGCTCAATGTAATCGTGAAGCATAAAGATTTTGATATACAT 468
OY 144 AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr 163
DB 469 GACATCATCTTCAATTTCTGCT-----ATTCATTAACACCAACC 504
OY 164 AspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGln--- 182
DB 505 GAACCTTACAAACAAAGTACTCGAAGAAATCGAGGAATGCTGCCCAATTTGTAACATA 564
OY 183 ---ThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
DB 565 ACAGATGCTTACGAAGAATTAGACAAGATTATCGACATTAACCTTAATAATATGCG 621

Search completed: June 12, 2003, 13:58:39
Job time : 1938 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 10:23:20 ; Search time 70 Seconds
(Without alignments)
1927.683 Million cell updates/sec

Title: US-09-777-566a-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPRPOSA.....IVNEARIPACSLRHHHHH 440

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US0977566/rnat_05062003_154804_6427/app.query.fasta_1.583
-DB=Issued_Patents_NA -OPMT=fastap -SOFTX=tni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cd1
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US0977566.ecgn.1.1.40.tunat.05062003_154804_6427 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -NAT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	3	US-09-259-214-1
2	2302	100.0	1323	4	US-09-318-528-1
3	2302	100.0	1323	4	US-09-291-931-1
4	2194.5	95.3	1272	2	US-08-910-798-1
5	352	15.3	5975	1	US-08-920-812-3
6	352	15.3	5975	1	US-08-920-827-23
7	352	15.3	5975	1	US-08-921-177-23
8	352	15.3	5975	1	US-08-362-57C-23
9	352	15.3	5975	2	US-08-920-828-23
10	119	5.2	1571	4	US-08-868-435-32
11	119	5.2	1571	4	US-08-744-231-32
12	119	5.2	1571	4	US-09-044-718-10

13	117	5.1	1588	2	US-09-146-283-1	Sequence 1, Appl
14	117	5.1	1588	3	US-08-579-823A-1	Sequence 1, Appl
15	117	5.1	1588	4	US-09-344-195-1	Sequence 1, Appl
16	117	5.1	3061	2	US-08-692-787-47	Sequence 47, Appl
17	117	5.1	3061	4	US-09-097-199-47	Sequence 47, Appl
18	112	4.9	11604	4	US-09-385-028-13	Sequence 13, Appl
19	112	4.9	15079	4	US-09-385-028-1	Sequence 1, Appl
20	111.5	4.8	80161	3	US-09-036-987A-1	Sequence 1, Appl
21	111.5	4.8	80161	4	US-09-370-700-1	Sequence 1, Appl
22	108	4.7	5153	5	PCT-US95-04910-8	Sequence 8, Appl
23	106.5	4.6	3147	4	US-08-887-534A-84	Sequence 84, Appl
24	105	4.6	1912	4	US-08-868-435-11	Sequence 11, Appl
25	105	4.6	1912	4	US-08-744-231-11	Sequence 11, Appl
26	104	4.5	1642	3	US-08-993-359-27	Sequence 27, Appl
27	103.5	4.5	2071	1	US-07-923-724-1	Sequence 1, Appl
28	103.5	4.5	2071	2	US-08-609-426A-1	Sequence 1, Appl
29	103.5	4.5	2071	2	US-08-374-652C-3	Sequence 3, Appl
30	100	4.3	12508	4	US-09-655-270A-1	Sequence 1, Appl
31	100	4.3	12523	4	US-09-651-941-1	Sequence 1, Appl
32	100	4.3	12523	4	US-09-955-597-1	Sequence 1, Appl
33	100	4.3	15378	3	US-08-785-420-1	Sequence 1, Appl
34	99.5	4.3	1845	4	US-08-868-435-30	Sequence 30, Appl
35	99.5	4.3	1845	4	US-08-744-231-30	Sequence 30, Appl
36	99.5	4.3	1845	4	US-09-044-718-7	Sequence 7, Appl
37	99	4.3	3836	2	US-08-216-260-1	Sequence 6, Appl
38	99	4.3	28958	1	US-08-258-261B-6	Sequence 6, Appl
39	99	4.3	28958	1	US-08-456-837-6	Sequence 6, Appl
40	99	4.3	28958	1	US-08-457-342-6	Sequence 6, Appl
41	99	4.3	28958	1	US-08-457-646A-6	Sequence 6, Appl
42	99	4.3	28958	1	US-08-458-076A-6	Sequence 6, Appl
43	99	4.3	28958	1	US-08-764-233A-4	Sequence 4, Appl
44	99	4.3	28958	1	US-08-457-335A-6	Sequence 6, Appl
45	99	4.3	28958	1	US-08-729-214-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 610719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: DIVER370-1
CURRENT APPLICATION NUMBER: US/09/259, 214A
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: 08/910, 798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
NAME/KEY: misc_feature
LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1
Alignment Scores:
Pred. No.: 3.03e-239
Score: 2302.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Gaps: 0
US-09-777-566a-2 (1-440) x US-09-259-214-1 (1-1323)

```

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCATTTTATCTCTTCGATTCGCTTAACCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTACCCGAGCTGAAGCTGGAAAGTGGTGATGTGATGCTGATGCTATGCT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCCAAACCAAGCCAGCCAGCACTGATGAGATGTCACCCGAGAGCATGGCCA 180
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCGTGGCGGTAAACTGGGTGGCTGACCGCGNGGTGGAGCTATATCCCTATCTTC 240
QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTTCCAAACGCGACGCTGCTGTACCCAGGATTTGCTGGCAAAAAGGCTGCGCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGTGTCAGTGTGCGATTATTGCTGATGTCGACGAGCGTACCCGTAACAGCGCA 360
QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGACCTGCTGCTGCAATACCCGTAACAGCGACATGACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCACATCCCGATCCCTTTATTCCTTAATCCTTAACAACTGGCGTTTGGCAACTGATACCG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCAGACGCGAGAGGGTCAATTGCTGACTTACCGCGCAT 540
QY 181 ArgGlnTrpAlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuGly 200
Db 541 CGGCAACAGGGGTTTGGCAACTGGAACTGGAGCGGTCTTAATTTTCCCAATCAAACTTGTGC 600
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAAGCTGAAGAAACAGAGCAAGAGCTTTCATTAACGACGCTTACCATCGGAATCTC 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACCGCGCAATGCTCATTAACCGGTGGGTAACCTGCAATCAATCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleGlyMetProGlyLysArgGlyLeuThr 260
Db 721 GAGATATTCTCTGCAACAAAGCACAGGAAATGCGGAGCGGGGTGGGAAAGATCACCC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 781 GATTACACACGAGTGAACACCTTGCTAGTTTGGCTAATACGGGCATTTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCAGAGTGTGGCCGACGCGCCACCCCGTATTGATGATCATGATGACAGCG 900
QY 301 LeuThrProHisSerProGluLysGlnIleArgValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCCATCCACCCCAAAAACAGCGGTATGGTGTGACATTAACCACTTCATCTAGT 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db 961 TTTATTTGCGGAGACAGATACACTGTGCAATCTCGGCGCGGCACTGAGACTCAACTGG 1020
QY 341 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuValPheGlnArgTrp 360
Db 1021 ACGCTTCCCGGTGACCGCGATACACGCGCGAGGTGTAACCTGCTTTGAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

```

```

Db 1081 CGTCGCTAAGGAAAGAACAGCCAGTACGATTCAGGTTTGGCTGCTTCAGACTTTACAG 1140
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGlyValLysLeuThr 400
Db 1141 CAGATGCGGTATTAACCCCGCTGCTATTAATTAATACCCCGCGAGAGTGAACCTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 GTGCGAGAGATGGAAGACGAAATGCGAGGCGATGTGTGTTGGCAGGTTTACGCAA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTAATGAAGACGAGATACCGGCTGGCACTTTGAGATCTCATCACCACCATCAG 1320

RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

Alignment Scores:
Pred. No.: 3,036-239 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-318-528-1 (1-1323)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAAGCGATCTTAATCCATTTTATCTCTTCGATTCGCTTAACCCGCAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTACCCGAGCTGAAGCTGGAAAGTGGTGATGTGATGCTGATGCTATGCT 120
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCCAAACCAAGCCAGCCAGCACTGATGAGATGTCACCCGAGAGCATGGCCA 180
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCGTGGCGGTAAACTGGGTGGCTGACCGCGNGGTGGAGCTATATCCCTATCTTC 240
QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100

```

```

Db      241  GGACATTACCAAGCCAGCGTCTGGTACCGACGAGATTGTCGGCAAAAAGGCGTCCCG 300
Qy      101  GlnserglyValAlaIleIleAlaaspValaspGluArgThrArgLysThrGlyGlu 120
Db      301  CAGCTGTGGTACGTCGGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGGCCAA 360
Qy      121  AlaPheAlaAlaGlyLeuAlaProaspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      361  GCCTTCGGCGCGCGGCTGGACACTGACTGTCATTAACCGTAACATACCAGACAGACATCG 420
Qy      141  SerSerProaspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      421  TTCAGTCCCGGATCGGTATTTAAATCCTTAAACCTGGCGGTGGCAACTGGATTAACCG 480
Qy      161  AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      481  AACGTGACTGACGGGATCTCTACGACGAGGCGAGGATTAATGCTGACTTTAACCGGCAT 540
Qy      181  ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGys 200
Db      541  CGGCAAAACGGCGTTCGGCAACTGGAACGGGTGCTTAATTTCCGCATCAAACTTGTGC 600
Qy      201  LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      601  CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAAACGAGGCAATTCGGAACCTC 660
Qy      221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      661  AAGGTGAGCGCCGCAACATGCTCATTAACCGGTGGGTAACGCTTGCATGCTGTACG 720
Qy      241  GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyLysArgIleThr 260
Db      721  GAGATATTTCTCTGCAACAGACAGAGGATGCCGAGCGGGGTGGGGAAGATCAACC 780
Qy      261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGln 280
Db      781  GATTCAACACAGTGGAAACCTTCTGATTTGCAATAACGGCAATTTATTTGCTACAA 840
Qy      281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db      841  CGCACCGCAGAGTGGCCCGACCGCGCCACCCCGTTATGATTTATCATGTCACGCG 900
Qy      301  LeuThrProHisProProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
Db      901  TTGACGCCCCATCCACCGCAAAACAGCGTATGGTGTGACATTAACCACTTCACTACG 960
Qy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      961  TTTATTTGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Qy      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1021  ACGCTTCCCGGTACCGCGGATTAACAGCCCGCAGGTGTGAACTGTGTTTGAACGCTGG 1080
Qy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081  CGTGGGCTTAACGATTAACAGCAGATGATCAGGTTTCTGTCGTCCTTCACAGCTTTACG 1140
Qy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1141  CAGATGCGCTGATTAACCGCCGCTGCTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Qy      401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1201  CTGGCAGGATGTGAAGACGAAATGCGCAGGCGATGTGTTGTTGGCAGGTTTTCAGCAA 1260
Qy      421  IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db      1261  ATCGTGAATGAAGACACGATACCGCGCGTGCAGTTTGAATGTCATCAACATCACCATCAC 1320

```

RESULT 3
 US-09-291-931-1
 ; Sequence 1, Application US/09291931A

```

; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Alignment Scores:
Pred. No.: 3,03e-239 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-291-931-1 (1-1323)
Qy      1  MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      1  ATGAAGCGAGCTTATATCCATTATTTATCTTCTGATTCGCTTAACCCCGCAATCTGCA 60
Qy      21  PheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArgHisGly 40
Db      61  TTGCGTCAAGTGAACCGGAGCTGAAGCTGGAAGTGTGTGATTTGTGATGTCATATGCT 120
Qy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      121  GTGGCTGCTCCAAACCAAGGCCGCACTGATGACGATGTACCCCGCAGCGCATGGCCA 180
Qy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db      181  ACGTGGCGGTFMAAAGTGGTGTGCTGACACCGCGNGGTGGTGAAGCTATATGCTATCTC 240
Qy      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlySerPro 100
Db      241  GGACATTTACCAACGGCCACCGTGTGTACCGCGAGATTGTGTGGCAAAAAGGCGTCCCG 300
Qy      101  GlnSerGlyGlnValAlaIleIleAlaaspValaspGluArgThrArgLysThrGlyGlu 120
Db      301  CAGTGTGGTACAGTGGCATTTATGCTGATGTCGACGAGCGTACCCGTAACAGGCCAA 360
Qy      121  AlaPheAlaAlaGlyLeuAlaProaspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      361  GCCTTCGGCGCGCGGCTGGACACTGACTGATTAACCGTAACCGTAACCCAGCAGATACG 420
Qy      141  SerSerProaspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      421  TTCAGTCCCGGATCGGTATTTAAATCCTTAAACCTGGCGGTGGCAACTGGATTAACCG 480
Qy      161  AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      481  AACGTGACTGACGGGATCTCTACGACGAGGCGAGGATTAATGCTGACTTTAACCGGCAT 540
Qy      181  ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGys 200

```

```

Db      541 CGGCAACGGCGTTTCGGCAACTGGAACTGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Qy      201 LeuysargGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      601 CTTAAAGCTGAGAAACAGACGAAAGCTGTCTTAAACGACGACATTCATCCGGAATC 660
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      661 AAGGTAGCGCGCAATGTCTCAATTAACCGGTGGTACCCCTGCATCAATCGTGACG 720
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIlePhe 260
Db      721 GAGATATTTCTCTGCAACAAACAGACAGGGAATGCCGAGCGGGGTGGGAAAGATCAC 780
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db      781 GATTACACACGATGGAAACACTTGTAGTTGTGATACGGGCAATTTATTTGCTACAA 840
Qy      281 ArgThrProGluValAlaIleArgSerArgAlaThrProLeuLeuAspLeuIleMetAla 300
Db      841 CGCAGCGCAGAGGTGGCCGCGACGCCGCCACCCCGTTATTTGATGATGATGACGACG 900
Qy      301 LeuThrProHisProGlnLysGlnAlaIleArgValAlaThrLeuProThrSerValLeu 320
Db      901 TTGACGCCCATCCACCCGCAAAACAGCGGATATGATGATGATGATGATGATGATGATG 960
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      961 TTTATTTGCGGACACGATTAATCTGTGCAATCTCGCGGGCGACCTGAGCTCAACTCG 1020
Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1021 ACGCTTCCCGTCAGCCGCGATACACAGCGCGGAGGTGTAACCTGTGTGAACGCTGG 1080
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CGTGGCGTAACGATACACGACGATGATGATGATGATGATGATGATGATGATGATGATG 1140
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1141 CAGATGCGTGAATAAACCGCCCTGTCAATTAATACCGCGCGGAGAGTGAACACTGACC 1200
Qy      401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGGCAGGATGTGAAGACGAAATGCGAGGCGCATGTGCTGTGGCAGGTTTACGAA 1260
Qy      421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db      1261 ATCGTGAATGAAGACGATACCGCGCTGCAAGTTGAGATCTCATACCATCACCATCAC 1320

RESULT 4
US-08-910-798-1
: Sequence 1, Application US/08910798
: Patent No. 5876997
: GENERAL INFORMATION:
: APPLICANT: KRETEZ
: TITLE OF INVENTION: NOVEL PHYTASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/910,798
: FILING DATE: August 13,1997

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HALL, PH.D., LISA A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09010/029001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5099
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1272 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: PHYTASE
: FEATURE:
: NAME/KEY:
: LOCATION:
: US-08-910-798-1

Alignment Scores:
Pred. No.: 1,199-227 Length: 1272
Score: 2194.50 Matches: 423
Percent Similarity: 96.14% Conservative: 0
Best Local Similarity: 96.14% Mismatches: 0
Query Match: 95.33% Indels: 17
DB: Gaps: 1

US-09-777-566a-2 (1-440) x US-08-910-798-1 (1-1272)
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      1 ATGAAAGGAGCTTATATCCATTTTATCTCTTGATTCGCTGTAACCCGCAATCTGCA 60
Qy      21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValAlaIleValSerArgHisGly 40
Db      61 TTCGCTCAGAGTACCGCGGACCTGAAGCTGGAAGTGTGTGATGTCATGTCATGCT 120
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAlaThrProAspAlaTrpPro 60
Db      121 GTGCGTGTCCACCAAGGCGACGCACTGATGCAAGATGTCACCCGACGACATGGCCA 180
Qy      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db      181 ACCTGCGCGTAATAACTGGTTGGCTGACACCGCGGTGGAGCTAATGCGCTATCTC 240
Qy      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db      241 GGACATTACCAACGCGACGCTGTGTAGCCGACGAGATTGTGGCGAAAAGGCGTGGCCG 300
Qy      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgLysThrArgLysGlyLys 120
Db      301 CAGTGTGTCAGTGTCGATTAATTCGTATGTCGACGAGCGATACCCGTAACAGAGC--- 357
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      358 -----CAAGCAGATACG 369
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      370 TCCAGTCCCGATCCGATTTAATCTCTTAATAAAGTGGCTTGGCAACTGGATACCGC 429
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      430 AACGTGATGACGCGATCTCAGCAGGCGAGAGGCTCAATTTGCTGATTTACCGGGAT 489
Qy      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db      490 CGGCAACGCGGTTTCGGGAACGTGAAGCGGGTCTAATTTTCGCAATCAAACTTGTGC 549
Qy      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

```

```

|||||
Db 550 CTTAAACGTGAGAAACAGACGAAACCTGTCATTAAACGAGCATACCATCGAATC 609
OY 221 LysValSerAlaAspAsnValSerLeuThrGlyValValSerLeuAlaSerMetLeuThr 240
Db 610 AAGGTGAGGCGGACATATGTCATTAAACGGGGGTAAGCCCTCCATCAATGAGTACG 669
OY 241 GluLeuPheLeuLeuGlnGlnAlaGlnGlyMetProGlyIleProGlyIleThr 260
Db 670 GAGATATTCTCTCTGCAACAAGACAGGGAATCCCGAGCGGGGGTGGGAAGATCACC 729
OY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 730 GATTCACACACAGTGAACAACCTGTCAAGTTTCATACGCGCAATTTATTGTCACAA 789
OY 281 ArgThrProGlyValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
Db 790 CCGACCGCCAGAGTTCCCGCAGCCGCCGCCCTGTATTGATTTGATCATGGCAGCG 849
OY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 850 TTGACCGCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTACCCACTTACATCTG 909
OY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp 340
Db 910 TTATATGCCGACACAGATCTAATCTGGCAATCTGGCGGCGCAGTGGAGCTCACTG 969
OY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluIleThr 360
Db 970 ACGCTTCCCGGTACGCGGATTAACACGCGCCAGGTGTGAACCTGTGTTGTAACCTCG 1029
OY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnTrpLeuGln 380
Db 1030 CCGCGGCTAAGCATTAACAGCCAGTGCATTCAGGTTTCGCTGCTTCCAGACTTTACG 1089
OY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1090 CAGATCGTGATTAACAGCGCGCTGTATTAAATACGCCGCCGAGAGGTGAACATGACC 1149
OY 401 LeuAlaGlyCysGlnGluIleArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1150 CTGGCGGATGTGAAGAGCAATGGCAGGGCATGTGTCGTGGCAGAGTTTAAAGCA 1209
OY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1210 ATCGTGAATGAACAGCATACCGGCGTGCAGTTGAGATCTCATCACCATCAC 1269

```

RESULT 5

```

US-08-920-812-23
; Sequence 23, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Ede, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORGANISM: Klebsiella pneumoniae
; STRAIN: Clinical isolate KI-50
; US-08-920-812-23

Alignment Scores:
Pred. No.: 2,366-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 1 Gaps: 14

US-09-777-566a-2 (1-440) x US-08-920-812-23 (1-5975)
OY 2 LysAlaIleLeuIlePheProLeuSerLeuIleProLeuThrPro---GlnSerAla 20
Db 578 CAGGGGTGTACGCTGTTATTCCTGCGCGCTGCGCTGCGTGGCGGCTGCAATCTGCC 637
OY 21 PheAlaGlnSerGlnProGlyLeuLysLeuGlnSerValValIleValSerAlaGly 40
Db 638 GCCGCC-----GCGGACTGCGACGTGAGAAAGTGTGTCAGCTCACGCCGCCACG 668
OY 41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaThr 59
Db 689 ATTCGTCGCCGACAGCGCGGCAACCGGAAGCCATCAGCGCCGCCACGCCGCAACG 748
OY 60 ProThrTrpProValLysLeuGlyTyrPheLeuThrProArgGlyGlyLeuIleAlaTyr 79
Db 749 ACCGAGTGCACACCCCATGACGGGAGAGCTCACCGGCGCATGCGCGCGGTGTCAC 808
OY 80 LeuGlyHisTyrGlnArgGlnArgGlyLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
Db 809 AAAGGGCGTGGAGAGGCGCATTTACCGCCAGCTCGCTGCTG---CAGGCCGAGTGC 865
OY 100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGly 119
Db 866 CCAGCGGCGGATGATATACGTGCGCGCCAGCCCGCTGCGAGCGGACGACGACCGCC 925
OY 120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db 926 CAGGCGCTGTGATGATGCTCCCGCGCTGCGGCGCTGCTATTCATTTATGTC----- 979
OY 140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValLysGlnLeuAspAsn 159
Db 980 AGCGGGGATGCCATCCCTGTTTCAGACCGACAAAGTTCGCCGCCACGCAAAACGCC 1039
OY 160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGly 179
Db 1040 GCCCGGCA-GCTGCGCGGTGAAGAAGAGAGCGCGG-----GATCTGCC-GCA 1085
OY 180 HisArgGlnThrAlaPheArgGlyLeuGlnArgValLeuAsnPheProGlnSerAsnLeu 199
Db 1086 GGTGCGCAGCGC-----CTGGCGCCGACCATTCAGCTATTGAACAGCGCGTT 1133

```


QY 180 HsArgGlnThrAlaPheArgGlnLeuGlnAlaArgValLeuAsnDheProGlnSerAsnLeu 199
Db 1086 GGTGGCGCAGCGC-----CTGGCGCCACCATTCAGCTATGTGAACAGCGGTT 1133
QY 200 CysLeuLeuArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlu 219
Db 1134 TGT-----CAGGCCGATTAAGCCCTGCGCCGATTCGATACCC--CCGTGGCAG 1178
QY 220 LeuLysValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAlaSerMetLeu 239
Db 1179 GTGGACAGACGAGAAAGTGGGAGAGACCATTAAGCGGATGAGCCGATGAGCCCAATATG 1238
QY 240 ThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlyTyrGly 257
Db 1239 GTGGACAGCGCTGCTGCTGCTGAGTGAACCTGCTCTCAGCCAGCTGGCGTGGCGC 1298
QY 258 ArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyr 277
Db 1299 AAGATACCCAGGCGCAGGAGATACCGCCCTGCTGCTGCTTAACGAGAAACATACGAT 1358
QY 278 LeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
Db 1359 CTGAGTACGATGTGTGTATACCGCCGCAAAACCGCGGTGCTGCTCAACGCTATG 1418
QY 298 MetAlaAlaLeuThrProHisProProGlnLysGlnAlaTyrGlnAlaThrLeuProThr 317
Db 1419 CTCGACGCGCTCAACCGGAGCAGATCGAACGT-----ACGCTGGCTGCT 1463
QY 318 SerValLeuPheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGln 337
Db 1464 GCT-----GCTGGCATGTGACACCAATATGCTGCTGCGCACGCTGATGATAC 1511
QY 338 LeuAsnTrpThrLeuProGly---GlnProAspAsnThrProProGlyGlyLeuVal 356
Db 1512 TTATGCTGGACGCTGCGGCGCTACACCGGGGAAATATCCCGCGGACACAGCTGTGTG 1571
QY 357 PheGlnArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
Db 1572 CTGGACGCGCTGCGCAACGCGAGACGCGAGACGATATCGCGGCTATTTCACAGCC 1631
QY 377 GlnThrLeuGlnIleMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
Db 1632 CAGGCGCTCGACAGCTGCTGCTGCTGCGACAGCCCGGACGCGCATGCTGCTG 1691
QY 397 ValLysLeuThrLeuAlaGlnCysGlnGluArgAsnAlaGlnGlyMetCys----- 413
Db 1692 CAGGATGGCATAGCGCGGCTGCTGCTGCGACAGCCGCGCATGCTGCTGCTGCTG 1751
QY 414 ---SerLeuAlaGlnPheThrGlnIleValAsnGlnAlaArgIleProAla 429
Db 1752 GCGCGATTAACGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1802

RESULT 8
US-08-362-577C-23
Sequence 23, Application US/08362577C
Patent No. 5807673
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,577C
FILING DATE: 27-MAR-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Klebsiella pneumoniae
STRAIN: Clinical isolate KI-50
US-08-362-577C-23

Alignment Scores:
Pred. No.: 2,36e-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 1 Gaps: 14

US-09-777-566a-2 (1-440) x US-08-362-577C-23 (1-5975)

QY 2 LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrPro---GlnSerAla 20
Db 578 CAGGGGCTGTATGACCTGTTATTCGCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCG 637
QY 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValValIleAlaSerArgHisGly 40
Db 638 GCGCGC-----GCGGACTGCGACCTGAGAAATGCTGCTGACGCTACGCCACCGGT 668
QY 41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaTrp 59
Db 689 ATTGCTCCGCGGACGCGCGGCAACCGGAAAGCATGAGCGCCGACCGCCGACCGTGG 748
QY 60 ProThrTrpProValLysLeuGlnTyrTrpLeuThrProArgGlyGlyLeuIleAlaTyr 79
Db 749 ACCGAGTGGACCAACCCATGAGGGGAGAGCTCACCGGCATATGCGCGGTGTCAC 808
QY 80 LeuGlnHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCys 99
Db 809 AAGGCGCTGGGAGCGCAGCATTAACCGCCAGCTGCGCTGCTG---CAGCGCGGATGC 865
QY 100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspLysLysArgLysThrGly 119
Db 866 CCGACGCGGAGTGCATATACGTCGCGCCGACCGCGCTGACGAGCGGACGCGACCGCC 925
QY 120 GlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db 926 CAGGCGCTGTGATGCGCGCTTCCCGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCT 979
QY 140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsn 159
Db 980 AGCGGGATGCCATCCCTGTTTCAGACCGCAGATGTGCGCCGACCAACCGACCGCC 1039
QY 160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGly 179
Db 1040 GCCCGCA-GCTGGCGGCTGAAGAGAGAGAGCGCGG-----GATCTGCG-GCA 1085

```

OY 180 HisArgGlnThrAlaPheArgGluLeuGluArgValLeuAspPheProGlnSerAsnLeu 199
    |||||
Db 1086 GGTGGCGAGGCG-----CTGGCGCGGACCATTCACGATTTGAAACAGGCGGT 1133
OY 200 CysLeuYsArgGluYsGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlu 219
    |||||
Db 1134 TGT-----CAGCGCGATTAAGCCCTGCGATCTTCGATACC---CCGTGGCGAG 1178
OY 220 LeuYsValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAsnLeu 239
    |||||
Db 1179 GTCGAGCAGACCAAAAGTGGCAAGCACCATTTACGAGCTGAGCGTGCGCCATATG 1238
OY 240 ThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlyTyrGly 257
    |||||
Db 1239 GTGAGAGCGTGGCGCTGGCTGAGTGAATAACCTGCGCTTCGAGCAGCTGGCGGGCG 1298
OY 258 ArgIleThrAspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheYr 277
    |||||
Db 1299 AAGATCACCCAGGCGAGATCAACCCGCTGCGCTGTAACGAAACATACGAT 1358
OY 278 LeuGluGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
    |||||
Db 1359 CTGAGTACGATGTTTATTATACCGCGCAAAACCGGGGTGGCTGCTCAACGCTATG 1418
OY 298 MetAlaAlaLeuThrProHisProProGlnGlnAlaTyrGlyValThrLeuProThr 317
    |||||
Db 1419 CTCGACGCGCTCAAAACGAGGAGATCGAAGCT-----ACGCTGGCTGCT 1463
OY 318 SerValIleuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlu 337
    |||||
Db 1464 GCT-----GGTGGCATGACACACATATCGCATGCTGCGACGCGTGAATGAAC 1511
OY 338 LeuAsnThrThrLeuProGly---GlnProAspAsnThrProGlyGlyGluLeuVal 356
    |||||
Db 1512 TTTACTGTCGACCTCCCGGCTACACCGCGGGAATATCCCGCGGAGCGCTGGG 1571
OY 357 PheGluArgTyrArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPhe 376
    |||||
Db 1572 CTGGAGCGCTGGCGCAACGCGAAGAGCGAGAACCTATCTCGGGGTCTATTCCAGGCG 1631
OY 377 GlnThrLeuGlnGlnMetArgAspYsThrProLeuSerLeuAsnThrProProGlyGlu 396
    |||||
Db 1632 CAGGCGCTCGACGACCTCGCTGCTGCAAGACGCGGACGCGACCGCGATGCTGCGT 1691
OY 397 ValYsLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCys----- 413
    |||||
Db 1692 CAGAGGTGCATCAGCCGGGCTGCGTCAAGACCATTCGCGTGTGCTCCCTCCAG 1751
OY 414 ---SerLeuAlaGlyPheThrGlnIleValAsnGlnAlaArgIleProAla 429
    |||||
Db 1752 GCGGCTATTACCGCCCTCGGTCAAGGCTATCGACGATCCGCGCCGCGCG 1802

```

RESULT 9
US-08-920-828-23
Sequence 23, Application US/08920828
Patent No. 5853398

GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,828
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
;
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella pneumoniae
; STRAIN: Clinical Isolate KT-50
;
; US-08-920-828-23
;
; Alignment Scores:
; Pred. No.: 2,366-27 Length: 5975
; Score: 352.00 Matches: 116
; Percent Similarity: 43.25% Conservative: 73
; Best Local Similarity: 26.54% Mismatches: 211
; Query Match: 15.29% Indels: 39
; DB: 2 Gaps: 14
;
; US-09-777-566a-2 (1-440) x US-08-920-828-23 (1-5975)
;
OY 2 LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrPro---GlnSerAla 20
    |||||
Db 578 CAGGGCGCTGTACGCTGTATTATCGCTGCGCGCTGCGCTGCGCTGCAACTGCG 637
OY 21 PheAlaGlnSerGluProGluLeuYsLeuGlnSerValAlaIleValSerArgHisGly 40
    |||||
Db 638 GCCGCC-----GCGACTGCGCAGCTGGGAAAGTGTGAGCTCAGCGCCACAGGT 688
OY 41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaThr 59
    |||||
Db 689 ATTGCGTCGCGCAGCGCGCGCAACGCGGAGCATCGAGCGCGCCAGCGCGACCGTGG 748
OY 60 ProThrThrProValYsLeuGlyTyrPheThrProArgGlyGlyGluLeuIleAlaThr 79
    |||||
Db 749 ACCGAGTGGACCAACCATCAGCGGAGCTCACCGCGCATGCTATGCGCGCTGGTCAAC 808
OY 80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
    |||||
Db 809 AAAGGCGCTGGGAAAGCCAGCATTAACCGCCAGCTCGGCTGCTG---CAGCCGAGATC 865
OY 100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspValAspArgThrArgYsThrGly 119
    |||||
Db 866 CCGAGCGCGGATCGATATACGTGGCGCGCACCGCGCTGACAGCGGAGCGAGCGACCGCC 925
OY 120 GluAlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
    |||||
Db 926 CAGGCGCTGCTGATGAGCGCTTCCCGGCTGCGCGCTGCTGCTATCATATATGTC----- 979
OY 140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsn 159
    |||||
Db 980 ACGGCGGATCGGATCCCTGTTTCAAGACCGCAAGTGTGCGCCAGCAAGCAAGCAAGCC 1039
OY 160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGly 179
    |||||

```

```

Db 1040 GCCCGCCA-GCTGGCCGGTGAAGAGAGCCGGC-----GATCGGC-GCA 1085
Qy 180 HsArgGlnThrAlaPheArgGlnLeuGluArgValLeuAsnProGlnSerAsnLeu 199
Db 1086 GGTGGGAGGCGC-----CTGGGCGGACCATTCACGATTAAGAAAGCGCGCT 1133
Qy 200 CysLeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlu 219
Db 1134 TGT-----CAGGCGGATTAAGCCCTCCCATCTTCATACCC---CCGTGGCAG 1178
Qy 220 LeuLysValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAsnMetLeu 239
Db 1179 GTGACACAGACAAAAGTGGGAAGACCATTAAGCGGAGCTGAGCGGATGCGCAATATG 1238
Qy 240 ThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlnTyrGly 257
Db 1239 GTGAGACAGCTGCGTCCGCGCTGGAGTGAAGAACCTGCTTCAGCAGCAGCGGGGTGGGC 1298
Qy 258 ArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyr 277
Db 1299 AAGATACCCAGGCGAGGACGATACCGCCCTGCTGCTTACGGAAGAACTACGAT 1358
Qy 278 LeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
Db 1359 CTGAGTACAGATGTTGTATACCGGCGCAAAAAGCGGGTGGTGGCTGCTCAACGCTATG 1418
Qy 298 MetAlaAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThr 317
Db 1419 CTCGACGCGGTCAACCGGAGCAATCGAACGT-----ACGCTGGCGCT 1463
Qy 318 SerValLeuPheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlu 337
Db 1464 GCT-----GCTGCCATGACACCAATATGCTGCTGGCGACCGCTGATGATAC 1511
Qy 338 LeuAsnTrpThrLeuProGly---GlnProAspAsnThrProProGlyGlyGlnLeuVal 356
Db 1512 TTAGCTGGCAGCTGCGGGCTACAGCCGGGGAATATCCCGCGGCGACACCTGGTG 1571
Qy 357 PheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
Db 1572 CTGAGCGCTGGCGGCAACGCGAAGACGCGAGAACGCTATCTGGGGCTATTTCAGGCC 1631
Qy 377 GlnThrLeuGlnGlnMetArgAspLysTrpProLeuSerLeuAsnThrProProGlyGlu 396
Db 1632 CAGGGCTCGACAGACCTGCTGCTGCAAGACCGCGGACGCGGACGATGCTGCGT 1691
Qy 397 ValLeuLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCys----- 413
Db 1692 CAGAGTGGCATACGCGGGCTGCTCAGACCGGATGCTGCTGCTGCTGCTGCTGCTGCT 1751
Qy 414 ---SerLeuAlaGlyPheThrGlnIleValAsnGlnAlaArgIleProAla 429
Db 1752 GCGGCTATTACCGCCCTCGTCAATGACGATACGATACGCGCGCGG 1802

```

RESULT 10

```

US-08-868-435-32
Sequence 32, Application US/08868435
Patent No. 6291221
GENERAL INFORMATION:
APPLICANT: Van Loon, Adolphus
APPLICANT: Mitchell, David
TITLE OF INVENTION: POLYPEPTIDES WITH PHYLASE ACTIVITY
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/868,435
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/744,231
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Kass, Alan P
? REGISTRATION NUMBER: 32142
? REFERENCE/DOCKET NUMBER: Case Docket 9339
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (201) 235-4205
? TELEFAX: (201) 235-2363
? INFORMATION FOR SEQ. ID NO: 32:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1571 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
? HYPOTHETICAL: NO
? ANTI-SENSE: NO
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(43..89, 147..1494)
? FEATURE:
? NAME/KEY: Intron
? LOCATION: 90..146
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 894..999
? OTHER INFORMATION: /note= "Position of PCR fragment"
US-08-868-435-32

Alignment Scores:
Pred. No.: 0.00419 Length: 1571
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.17% Indels: 114
Gaps: 24

US-09-777-566a-2 (1-440) x US-08-868-435-32 (1-1571)
Qy 7 ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro 26
Db 253 CCATCTCTTTCGCTCGAGAGCAGCTGCTCGGTGAGTAAAGCTTCCCAAG----- 303
Qy 27 GlnLeuLysLeuLysSerValIleValSerArgHisGlyValArgAlaProThr--- 45
Db 304 GATTGCCGATCACCTTGTGACAGGTGCTATCGCGCCATGAGACGCGGTACCCACACG 363
Qy 46 -----LysAlaThrGlnLeuMetGlnAspValThrProAspAla----- 58
Db 364 TCCAGAGCAAAAGATAGAGCTTGTGACGCGCATCCAGCCAAATGACCCAGCTTC 423
Qy 59 -----TrpProThrTrpProValLysLeuGlyTrp-----LeuThr 70
Db 424 AAGGCAAGTTGCTTTTGAAGAGCTATACCTATACCTGGGTGCGGATGACCTACT 483
Qy 71 ProArgGlyGlyLeuLeuIleAlaTyrLeuGly-----HisTyrGlnArgGlnArgLeu 88
Db 484 CCGTTGGGAGAGACAGCTGCTGAACCTCGGCGATCAAGTTCACCAAGAGTCAAGGCT 543
Qy 89 ValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIle 108
Db 544 CTGGCGGCGAGTGTGGCGC-----TTTATTCCG 573
Qy 109 AlaAspValAspGluArgThrArgLysTrpGlyGlnAlaPheAlaIleGlyLeu----- 126

```

```

Db 574 GCGTCAGGCTCGGACCGGGTTATGCTTCGGAGAGAACTTCATCGAGGGGTTCCAGCAG 633
QY 127 -----AlaProaspCysAlaIleThrVal 134
Db 634 GCGAGGTGGCTGATCCTGGCCGACGACACCGCGCGCTCCGGCGATTAGTGTGATTAT 693
QY 135 HisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuThrGlyVal 154
Db 694 ---CCGAGAGCGAGACGTTCAAC-----AATACGCTGGACACAGCGTGTG 735
QY 155 CysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIle 174
Db 736 TGC-----ACGAGATTGAGCGGACGATGCTGCGAGATGAGAGTT 774
QY 175 ---AlaaspPheThrGlyHisArgGlnThrAlaPheArgGlnLeuGlnArgValIleuAsn 193
Db 775 GCGGCCAATTTCAC-----GCGCTCTTT 798
QY 194 PheProGlnSerAsnLeuCysLeuysArgGlnGlnAspGlnSerCysSerLeuThr 213
Db 799 GCACCC-----GACATCGAGCTCGCGCGGAGAGACATCTTCGCGGTGACGCGACGACA 852
QY 214 GlnAlaLeuProSerGlnLeu-----LysValSerAlaAspAsnValSerLeuThrGly 231
Db 853 GACGAGGACGTTGTCACTTAATGACATGCTGCTTGATGACGTTACGCGGACGCGACAC 912
QY 232 AlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnIleGlnIleMet 251
Db 913 GACGAGACGCTGCTACCGCTTGTCTGCTCACTCTTCACTACACATGAG----- 960
QY 252 ProGlnProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuSerLeu 271
Db 961 -----TGG-----AAGAGTACACACTACTTGCATGCTTG 990
QY 272 HisAsnAlaGlnPheThrLeuLeuGlnIleArgThrProGlnValAlaArgSerArgAlaThr 291
Db 991 -----GGCAGACTACTACGCTACGCGGAGGACCAACCTCTGGGACCGGCTCAGGGGATA 1044
QY 292 ProLeuLeuAspLeuIleMetAlaIleuThrProHisProGlnIleGln----- 309
Db 1045 GGGTTCACCCAGACGCTGATGCGCGGTGACTGCTTCCGCGAGGACGACACACACAC 1104
QY 310 -----AlaTrpGlyValThrLeuProThrSerValLeuPheIleAla 323
Db 1105 ACTACTGACACTAGTCTCCACACCGCGCACCTTCCCTTGAACGCTACATGATACGTC 1164
QY 324 -----GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrpThr 341
Db 1165 GACTTTTTCACAGCAACAGCAATGGTTTCCATCTTCTTGCATTG----- 1209
QY 342 LeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrpArg 361
Db 1210 -----GGCCTGTACACGCGACTGACCTTGTCCGCGACCTCGGTGGAAGCGCCAG 1263
QY 362 ArgLeuSerAsp---AsnSerGlnTrpIle-----GlnValSerLeuValPheGln 377
Db 1264 GAATTGGATGGATTATCTGATCCTGGGTGGTGGCTTCCTGGCGCGGACCTCACTTCCGAG 1323
QY 378 ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnVal 397
Db 1324 ACGATGCAATGCAATCGGAAAGAGCGCTTT---GTTGCGGCTTGTGATTATGACCGG 1380
QY 398 LysLeuThrLeuAlaGlyCysGlnGlnIleArgAsnAlaGlnIleMetCysSerLeuAlaGly 417
Db 1381 GTTGTGCCACTGCAATGCGGCGATGTG---GACAGCTGGGCGATGCAAGCTGATGAC 1437
QY 418 PheThrGlnIleValAlaAsnGlnAlaArg 426
Db 1438 TTTGTCAAGGATGAGTTGGCGCAGA 1464

```

RESULT 11
 US-08-744-231-32
 ; Sequence 32, Application US/08744231

```

; Patent No. 6358722
; GENERAL INFORMATION:
; APPLICANT: Van Loon, Adolphus
; APPLICANT: Mitchell, David
; TITLE OF INVENTION: POLYPEPTIDES WITH PHYTASE ACTIVITY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/744,231
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,757
; FILING DATE: 18-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: Case Docket 9339
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ. ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(43..89, 147..1494)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 90..146
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 894..999
; OTHER INFORMATION: /note= "Position of PCR fragment"
; US-08-744-231-32

Alignment Scores:
Pred. No.: 0.00419 Length: 1571
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.17% Indels: 114
DB: 4 Gaps: 24

US-09-777-566a-2 (1-440) x US-08-744-231-32 (1-1571)
QY 7 ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGlnPro 26
Db 253 CCATCTTTTCGCTCGAGGACGAGCTGCTCGCTGCGATGAGCTTCCCAAG----- 303
QY 27 GlnLeuLysLeuGlnLysValValIleValSerArgHisGlyValArgAlaProThr----- 45
Db 304 GATTGCCGATCACCTTGATACAGGTGATGCGCGCATGAGCGCGGATACCAACACAC 363
QY 46 -----LysAlaThrGlnLeuMetGlnAspValThrProAspAla----- 58

```



```

Db 156 AGTCCACT-----GACACCTTTCCACATGAC 182
Oy 62 -----TriProValLysLeuGlyTriPleuThProArgLysGlyLyu 75
Db 183 CCCATAAAGAAATCCTCATGGCCACAGAGATTGGCCAACTACACCAGCTGGCATGGAG 242
Oy 76 LeuIleAlaTyRLeuGlyHisTyRGlInArgGlInArgLeuValAlaAspLysLeuAla 95
Db 243 CAGCATTTAGAACTTGAGAGTATATAGAAAGAGATAT----- 281
Oy 96 LysLysGlyCysProGlnSer-----GlyGlInValAlaIleIleAlaLysValAsp 112
Db 282 AAAAAATTCCTTGAAATGAGCTCATTAACAATGAACATGAACGTTTATTCGAGACAGACGTT 341
Oy 113 GluArgThrArgLysThrGlyLysAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle 132
Db 342 GACCGGACTTTGATGAGTCCCTATGACAAACCTGGCAGCCCTTTTCCCCACAGAGTGC 401
Oy 133 ThrValHisThrGlInAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db 402 AGCATCTGG-----AATCCTATTCCTACTGCGCAGCCCATCCCGGTG 443
Oy 153 GlyValCysGlInLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGly 172
Db 444 CACACAGTTCCTCTTCTGAA----- 464
Oy 173 SerIleAlaAspPheThrGlyHisArgInThrAlaPheArgLysLeuGlyLysValLeu 192
Db 465 -----CATCAGTTGCTA 476
Oy 193 AsnPheProGlnSerAsnLeuLysLysArgLysGlyLysGlyAspLysSerLysLeu 212
Db 477 TACCTGACCTTCAGAAC---TGCCCTCGCTTTCAAGAACTTGAGAGTGAAGTTCGAA 533
Oy 213 ThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspAsnValSerLeuThrGlyAla 232
Db 534 TCAGAGGAATTCACAGAGAGCTGCACCTTTAAGGATTTTATAGCTTACCTTGGGAAA 593
Oy 233 ValSer-----LeuAlaSerMetLeuThrGlyLysLeuPhe 243
Db 594 CTTTCAGATATCATGAGCCAGGACCTTTTGGAAATTTGGAGTAAGTTCACAGCCCTTAA 653
Oy 244 LeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTriPlyArgLysThrAspSerHis 263
Db 654 TATTGTGACAGTTCACAAATTTCACTTACCTTCCTGGCCACATGAGACACCATGACT 713
Oy 264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGlnArgThrPro 283
Db 714 AGTTGACAGAAATTTGCAGAAATTTGCCCTGCTGCTCTATGAGAAATTCACAGCAGAAA 773
Oy 284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrPro 303
Db 774 GAGAAATCTAGGCTCCAGAGGGGTGTCCTGTCATGAATCCTCAATCAGATG----- 827
Oy 304 HisProProGlnLysGlnAlaTyRGlYValThrLeuProThSerValLeuPheIleAla 323
Db 828 -----AAGAGACCAACTCAGATACCAAGCTACAAAACCTTATCATGTATCT 875
Oy 324 GlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlnLeu---AsnTrpThrLeu 342
Db 876 GCGCATGACACTACTGTGAGTGCCCTACAGATGCGCTAGATGTTTACACAGCACTCTT 935
Oy 343 ProGlyGlnProAspAsnThrProProGlyGlyLysLeuValPheGluArgTrpArgArg 362
Db 936 CCT-----CCCTATGCTTCTTGCCACTTGACGGAATTTACTTTAGAAAGGGGAGTAC 989
Oy 363 LeuSerAspAsnSerGlnTriPleuGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Db 990 TTT-----GTGAGATGATCTACTATCGAAT----- 1013
Oy 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyLysLeuThrLeuAla 402
Db 1014 -----GAGACGACAGCAGACCGCTATCCCTCATGTACCT 1049

```

```

Oy 403 GLYCYSGLUGLUARGASNALAGLNGLYMETCYSSERLEUALAGLYPHERTHRGLINLEVAL 422
Db 1050 GGCCTGAGCCCT-----ACCTGCTCCTGAGAGGTTTCTGAGCTGGTT 1094

RESULT 14
US-08-579-823A-1
; Sequence 1, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; APPLICANT: Wu, Hongyu
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Delinger & Associates
; STREET: 350 Cambridge Ave, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 336
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF
; IMMEDIATE SOURCE:
; LIBRARY: prostate carcinoma Lncap.FGC; PBMC

US-08-579-823A-1

Alignment Scores:
Pred. No.: 0.00702 Length: 1588
Score: 117.00 Matches: 87
Percent Similarity: 34.09% Conservative: 63
Best Local Similarity: 19.77% Mismatches: 178
Query Match: 5.08% Indels: 112
DB: 3 Gaps: 15

US-09-777-566A-2 (1-440) x US-08-579-823A-1 (1-1588)

Oy 3 AlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAla 22
Db 51 AGCCTTAGCCTTGCGCTTCTGCTTTTCTGCTAGACCGAGGTGTACTAGCC 110
Oy 23 GlnSerGluProGlnLeuLysLeuGlnSerValAlaIleValSerArgHisGlyValArg 42
Db 111 AAG-----GAGTTGAGATTGTGACTTTGTGTCTTGGCATGGAGACCA 155

```

```
Qy 43 AlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThr--- 61
   ::::|
Db 156 AGTCCCAAT-----GACACTTTCCCACTGAC 182
Qy 62 -----TTrProValLysLeuGlyTTrPleuThrProArgLysGly 75
   ::::|
Db 183 CCCATAAGAAATCTCATGGCCAAAGATTGGCCAAACCCAGCCAGCTGGAG 242
Qy 76 LeuLeuAlaTyrLeuGlnLysTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAla 95
   |||||
Db 243 CAGCATTAATGACATTGAGAGATTATTAAGAAAGATAT----- 281
Qy 96 LysLysGlyCysProGlnSer-----GlyGlnValAlaLeuAlaAspValasp 112
   ::::|
Db 282 AGAAATATCTTGATGATGAGTCTTAATAACATGACAGCGTTTATATTCGAAGCACAACGCTT 341
Qy 113 GlnArgThrArgLysThrGlyGlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaLe 132
   ::::|
Db 342 GACCGGACTTGTGATGAGTGTATGACAAACCTGGACGCCCTGTTCCTCCCGAAGAGTGTTC 401
Qy 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
   ::::|
Db 402 AGCATCTGG-----ATCCATATCTAGCTGTGGACCCCATCCCGGTG 443
Qy 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaLeuLeuSerArgAlaGly 172
   |||
Db 444 CACACAGTCTCTCTTTGAA----- 464
Qy 173 SerLeuAlaAspPheThrGlyHisArgGlnThrAlaPheArgGlyLeuGlnArgValLeu 192
   ::::|
Db 465 -----GATCAGTGTGCTA 476
Qy 193 AsnPheProGlnSerAsnLeuLysArgGlyLysGlnAspGlnSerCysSerLeu 212
   |||||
Db 477 TACCTGCCCTTCAGAGAAC---TGCCCTCGTTTCAAGAACTTGAGAGAGACTTTGAAA 533
Qy 213 ThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspAsnValSerLeuThrGlyAla 232
   ::::|
Db 534 TCAGAGGAATTCAGAAAGAGGCTGCACCTTAATAAGATTTATATGACCTATGGGAAAA 593
Qy 233 ValSer-----LeuAlaSerMetLeuThrGluLeuPhe 243
   ::::|
Db 594 CTTTCAGGATTACATGGCCAGGACCTTTTGGAAATTGTGAGTAAGCTACAGACCTTTTA 653
Qy 244 LeuLeuGlnGlnAlaGlnGlyMetProGlyTTrPLeuArgLysThrAspSerHis 263
   ::::|
Db 654 TATGTGTAGAGTGTTCACAAATTTACTTTACCTCTGCGCCATGGAATTCACAGAGAAAA 713
Qy 264 GlnTTrPAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrPro 283
   ::::|
Db 714 AAGTTGAGAGAAATTGTCAGAATTTGCCCTGTCCCTATGGAATTCACAGAGAAAA 773
Qy 284 GlnValAlaArgSerArgAlaThrProLeuLeuAspLeuLysMetAlaAlaLeuThrPro 303
   ::::|
Db 774 GAGAAATCTAGCTCCCAAGGGGTGTCTGTCATGAATCCATCAATCACATG----- 827
Qy 304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheLeuAla 323
   |||||
Db 828 -----AAGAGGCAACTCAGATACCAACGCTCAAAAACTTATATCTGTATTCT 875
Qy 324 GlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlnLeu---AsnTTrPThrLeu 342
   |||||
Db 876 GCGCATGACACTACTGTAGTGGCCTACAGATGGCGCTAGATGTTTACACAGCACTCCTT 935
Qy 343 ProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGlnArgTTrPArgArg 362
   |||||
Db 936 CCT-----CCCTATGCTTCTTGGCCACTTGACGGAATTTACTTTGAGAAGGGGAGATAC 989
Qy 363 LeuSerAspAsnSerGlnTTrPLeuGlnValSerLeuValPheGlnThrLeuGlnLysMet 382
   |||||
Db 990 TTT-----GTGGAGATGTACTATCCGAAT----- 1013
Qy 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThrLeuAla 402
```

```
Db 1014 -----GAGACGACGACGAGCCGTATCCCTCATGCTACT 1049
Qy 403 GlyCysGlnGluArgAsnAlaGlnLysTyrCysSerLeuAlaGlyPheThrGlnLeuAl 422
   |||||
Db 1050 GGCTGCAGCCCT-----AGCTGTCCCTGTGAGAGGTTTGTGAGCTGTT 1094

RESULT 15
US-09-344-195-1
Sequence 1, Application US/09344195
Patent No. 6210662
GENERAL INFORMATION:
Applicant: Laus, Reiner
Mu, Hongyu
TITLE OF INVENTION: Immunostimulatory Compositions
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Dehlinger & Associates
STREET: 350 Cambridge Ave. Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/344,195
FILING DATE: 24-Jun-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/146,283
FILING DATE: 03-SEPT-1998
ATTORNEY/AGENT INFORMATION:
NAME: Judge, Linda R.
REGISTRATION NUMBER: 42,702
REFERENCE/DOCKET NUMBER: 7636-0010.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-0880
TELEFAX: 650-324-0960
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1588 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: linear
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF fusion gene; Fig. 1
IMMEDIATE SOURCE:
LIBRARY: prostate carcinoma LNCaP, PBC
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-344-195-1

Alignment Scores:
Pred. No.: 0.00702 Length: 1588
Score: 117.00 Matches: 87
Percent Similarity: 34.09% Conservative: 63
Best Local Similarity: 19.77% Mismatches: 178
Query Match: 5.08% Indels: 112
DB: 4 Gaps: 15

US-09-777-566a-2 (1-440) x US-09-344-195-1 (1-1588)
Qy 3 AlaLeuLeuPheLeuSerLeuLeuLeuProLeuThrProGlnSerAlaPheAla 22
   ::::|
```

Db 51 AGCCTAGGCTTGGCTTCTGTTCTTTTCTGCTAGACCGAAGTACTAGCC 110
QY 23 GlnSerGluProGlnLeuLysLeuGlnSerValAlaIleValSerArgHisGlyValArg 42
Db 111 AAG-----GAGTTGAGATTGTGACTTGTGGTGGTTCGGCATGGACCGCA 155
QY 43 AlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpProThr--- 61
Db 156 AGTCCCAT---GACACCTTTCCCACTGAC 182
QY 62 -----TrpProValLysLeuGlyTrpLeuThrProArgGlyGly 75
Db 183 CCCATAAAGGAATCCCATGCGCCACAGACGATTGGCCCACTCCACGCTGGGCAATGGAG 242
QY 76 LeuIleAlaTrpLeuGlnLysIleTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAla 95
Db 243 CAGCATTATGAACTTGGAGAGATATATAGAAAGAGATAT----- 281
QY 96 LysLysGlyCysProGlnSer-----GlyGlnValAlaIleIleAlaAspValAsp 112
Db 282 AGAAATTTCTTGATGATGCTCTATATAACATGAACAGCTTATATTCGAAGCACAGAGCTT 341
QY 113 GluArgThrArgLysThrGlyGlnAlaPheAlaIleGlyLeuAlaProAspCysAlaIle 132
Db 342 GACCGGACTTTGATGATGCTCTATGACAAACCTGGCAGCCCTGTTCCCCACAGAGGTTC 401
QY 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db 402 AACATCTGG-----AATCTATCTCTGCTGCGACGCCATCCGCGTG 443
QY 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGly 172
Db 444 CACACAGTTCTCTTTCTGAA----- 464
QY 173 SerIleAlaAspPheThrLysIleArgGlnThrAlaPheArgGluLeuGlnArgValLeu 192
Db 465 -----GATCAGTTGCTA 476
QY 193 AsnPheProGlnSerAsnLeuLysLeuLysArgGluLysGlnAspGluSerCysSerLeu 212
Db 477 TACCTGCCCTTCAGGAC---TGCCCTCGTTTCAGAACTTGAGAGTGAAGACTTTGAAA 533
QY 213 ThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAla 232
Db 534 TCAGAGGAATTCAGAAAGCTGCACCTTATAGATTTATAGCTACTTGGGAAA 593
QY 233 ValSer-----LeuAlaSerMetLeuThrGluIlePhe 243
Db 594 CTTTCAGATTCATGCGCCAGACCTTTTGGAAATTGGAGTAAAGTCTACGACCTTAA 653
QY 244 LeuLeuGlnGlnAlaGlnLysMetProGluTrpGlyArgIleThrAspSerHis 263
Db 654 TATTGGAGAGTTCACAAATTTCACCTTACCTCCCTGGCCACGTGAGACACCATGACT 713
QY 264 GlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlleuLeuGlnArgThrPro 283
Db 714 AAGTTGAGAGAAATTGTCAGAAATTGCTCCTGCTCCCTATGAGAAATTCACAGAGAAA 773
QY 284 GluValAlaArgSerArgLalaThrProLeuLeuAspLeuIleMetAlaIleAlaLeuThrPro 303
Db 774 GAGAAATCTAGGCTCCAGGGGGTGTCCGTCAATGAATCTCAATCAGATG----- 827
QY 304 HisProProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeuPheIleAla 323
Db 828 -----AAGAGAGCACTCAGATACCAAGTACAAAACCTTATCATGTATCT 875
QY 324 GlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeu---AsnTrpThrLeu 342
Db 876 GCGCATGACACTACTGAGTGGCTTACAGATGGCGCTAGATGTTTACAAACGACATCCTT 935
QY 343 ProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluArgTrpArgArg 362
Db 936 CCT-----CCTATGCTTCTTGGCACCTTGACGGAATGTACTTGGAGAGGGGAGGTAC 989

QY 363 LeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGlnLysMet 382
Db 990 TTT-----GTGAGATGTAATAAGAAAT----- 1013
QY 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThrLeuAla 402
Db 1014 -----GAGACGACACGAGCCGATCCCTCATGCTCACT 1049
QY 403 GlyCysGluGlnArgAsnAlaGlnLysMetCysSerLeuAlaGlyPheThrGlnIleVal 422
Db 1050 GCCTGCAAGCCT-----AGCTGTCTCTGGAGAGGTTTGTGTAAGCTGTT 1094

Search completed: June 12, 2003, 12:11:09
Job time : 124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 11:33:34 ; Search time 226 Seconds

(without alignments)
2724.763 Million cell updates/sec

Title: US-09-777-566A-2

Perfect score: 2302

Sequence: 1 MKAILIFLSILIPFQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 870385 segs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPFO.spool/US09777566/Runat_05062003.154806.6495/app_query.fasta_1.583
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09777566.ecgn_1.1.146 @runat_05062003.154806.6495
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCIT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCITUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2302	100.0	9	US-10-034-985-1
2	2302	100.0	1323	US-09-777-566A-1
3	2302	100.0	1323	US-09-866-379-1
4	2235	97.1	1901	US-09-866-379-7

Result No.	Score	Query Match length	ID	Description
5	2226	96.7	1901	US-09-866-379-9
6	2221	96.5	1489	US-10-266-041-9
7	2221	96.5	1901	US-09-866-379-6
8	2220	96.4	1901	US-09-866-379-5
9	960.5	41.7	1326	US-10-021-723A-11
10	938.5	40.8	1325	US-10-021-723A-3
11	918.5	39.9	1325	US-10-021-723A-9
12	760	33.0	1431	US-10-021-723A-1
13	476.5	20.7	1230	US-10-062-848-10
14	472.5	20.5	1266	US-10-021-723A-5
15	472.5	20.5	1398	US-09-803-454-19
16	472.5	20.5	1425	US-10-213-990-23
17	472.5	20.5	1455	US-10-213-990-22
18	472.5	20.5	1571	US-10-062-848-10
19	472.5	20.5	3061	US-09-974-546-47
20	472.5	20.5	1372	US-09-765-111A-13
21	472.5	20.5	2707	US-10-097-340-242
22	472.5	20.5	1758	US-09-895-793-979
23	472.5	20.5	1758	US-09-822-827-979
24	472.5	20.5	2625	US-09-765-111A-5
25	472.5	20.5	2711	US-09-765-111A-24
26	472.5	20.5	2406	US-09-895-793-980
27	472.5	20.5	2406	US-09-822-827-980
28	472.5	20.5	1431	US-09-738-626-1548
29	472.5	20.5	3309400	US-09-738-626-1
30	472.5	20.5	3147	US-09-741-669-216
31	472.5	20.5	3147	US-09-815-242-5965
32	472.5	20.5	2105	US-10-153-668-30
33	472.5	20.5	64	US-10-021-723A-23
34	472.5	20.5	3097	US-09-962-832-227
35	472.5	20.5	103.5	US-09-738-626-1
36	472.5	20.5	536165	US-09-939-964-1
37	472.5	20.5	1725	US-09-738-626-763
38	472.5	20.5	101	US-09-915-814-3
39	472.5	20.5	101	US-09-759-508B-1
40	472.5	20.5	1803	US-10-072-307-57
41	472.5	20.5	1971	US-09-974-300-322
42	472.5	20.5	99.5	US-10-062-848-7
43	472.5	20.5	1803	US-10-072-307-19
44	472.5	20.5	99	US-10-108-605-302
45	472.5	20.5	98.5	US-10-160-758-5

ALIGNMENTS

RESULT 1
US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Keith
; TITLE OF INVENTION: NOVEL PHRASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature

LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1

Alignment Scores:

Pred. No.:	4.1e-264	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-777-566a-2 (1-440) x US-10-034-985-1 (1-1323)

```

QY      1 MetLysAlaIleLeuIleProPheLeuSerLeuIleIleProLeuThrProGlnSerAla 20
DB      1 ATGAAAGCATCTTAATCCCATTTTATCTCTTCTATTCGGTTAACCCCGCATGTGCA 60
QY      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
DB      61 TTCGCTCAGAGTGAGCCGAGCTGAAGTGAAGTGTGATGTCTCAGTGTGATGCT 120
QY      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAlaThrProAspAlaTPro 60
DB      121 GTGCGTGTCTCAACCAAGCCAGCCAGCACTGATGCAGATGTCACCCAGACGATGGCA 180
QY      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
DB      181 ACCTGGCCGGTAAACCTGGTGTGGTGCACACCGCGGAGTGTGATGCTAATTCGCTATC 240
QY      81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB      241 GGACATTTACCAACGCGCAGCGTGTGTAAGCCAGGATTCCTGGCGAAAAAGGCTCCCG 300
QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGln 120
DB      301 CAGTGTGTGTCAGTGTGCGGATTTATGCTGATGTCAGACGACGATACCCGTAAACAGCGAA 360
QY      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB      361 GCGTTCGCGCGCGGGGTGCACCTGACTGCATTAATTAACGCTAATCCAGGAGATAG 420
QY      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
DB      421 TTCAGTCCCGATCCGTTATTAATCCTCAAAAACGGGCTTGGCCAACTGGTAAACGGG 480
QY      161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB      481 AACGTGACTGACGCGATCTCTCAGCAGGCGACGAGGCTCAATGCTGATTTACGGGCA 540
QY      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB      541 CGGCAAAACGGCGTTTCGCAACTGCAAGCGGTGCTTAATTTCCGCAATCAAACTGTGC 600
QY      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB      601 CTTAAACGTAAGAACAGGACGAAAGCTTTCAATTAACGACGATTTACCATTCGAAATC 660
QY      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB      661 AAGGTGAGCGCGGACAAATCTCATTTAACCGGTGCGGTAAAGCTCCCATCATGCTGAC 720
QY      241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpPheArgIleThr 260
DB      721 GAGATATTCTCTCTCAACACACAGGAGGATGCCGAGCGGGGTGGGAGGATGACAC 780
QY      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
DB      781 GATTTCACACCACTGGAACACCTTGTAACTTGCATTAACGCCCAATTTATTGTGCTACAA 840
QY      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB      841 CGCAGCGCCAGAGGTGGCCGCGACGCGCCGCCGCTTAATGTGATTTGATCATGGGACGG 900

```

```

QY      301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValIleThrLeuProThrSerValLeu 320
DB      901 TTACAGCCCATCCATCCACCGCAAAAACAGCGCTATGAGTGTGACATTTACCCATTCAGTAC 960
QY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
DB      961 TTTATTCGCGGACACGATTAATCTGGCAAAATCTGGCGGCGGCGACTGAGCTCAACTGG 1020
QY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB      1021 ACCGTTCCGCTAGCGGATTAACACGCGCGCGGTGTGAATGTTGTTGAACCTGG 1080
QY      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB      1081 CGTCGCGTAAAGCATTAACACGCGATGATTCAGGTTTCGTGCTTCCAGACTTTACAG 1140
QY      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
DB      1141 CAGATCGGTGATTAACGCGCGCTGCTATTAAATACCGCCCGGAGAGGTGAACCTGACC 1200
QY      401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB      1201 CTGGCAGGATGTAAAGAGGAAATGCGCAGGCGATGTGTTGTCAGAGTTTACGCAA 1260
QY      421 IleValAsnGluAlaArgGlyLeuProAlaCysSerLeuArgSerHisHisHisHis 440
DB      1261 ATCGTGAATGAACACGACGATACCGCGGTGAGTTTGATGATCATCACCATCACTAC 1320

```

RESULT 2

```

US-09-777-566a-1
; Sequence 1, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566a-1

```

Alignment Scores:

Pred. No.:	4.1e-264	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-777-566A-2 (1-440) x US-09-777-566A-1 (1-1323)

```

Qy 1 MetLysAlaIleuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCGGTAAACCCGGAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCCTCTGAGATGAGCCGCGAGCTGAAAGTGTGATGATGTCAAGTCGTCATGCT 120
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 121 GTGGGTGCTCCACACCAAGCCAGCAGCACTGATGAGATGTCACCCAGACGCAATGGCCA 180
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACTGGTGTGGCTGACACCGCGNGGTGGAGACTAATGCCATATCTC 240
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTAACCAACCCAGCAGCTGTGTTACCCAGCATTTGCGCAAAAAGGCGTGGCCG 300
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAAGTCCGATTAATTCCTATGTCACGAGCGTACCCTGTAACACAGCGCA 360
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGGCCGCGGGCTGGCACTGACGTGCATTAACCGTACATACCAGCAATACG 420
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TTCAGTCCCGATCCGATTTTAAATCCTTAAAAACTGCGTTTGCACATGATTAACGCG 480
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrLys 180
Db 481 AACGTGACTGACCGGATCTCAAGCAGGAGAGGGTGAATTCGTGACTTAAACCGGCACT 540
Qy 181 ArgGlnThrAlaPheArgGluLeuGlnuArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db 541 CGGCAAAAGCGGTTTCGGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTGTGCG 600
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAGAAACAGAGCAAGAACTGTTCAATTAACGAGCGCATTAACATCGAATC 660
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGAGCGCGCAATGTCCTAATTAACCGGTGGTGAAGCTGCAATCAATGCTGACG 720
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCTCTCCACAAACAGCAACAGGAATCCGCGAGCCGGGTGGGGAAGATCACC 780
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACACAGTGAGAACACCTTGCTAGTTGATTAACGCGCATTTATTTGCTACAA 840
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCCAGAGGTTGGCCGAGCGCGCACCCCGTAAATGATTTGATTCATGTCAGCG 900
Qy 301 LeuThrProHisProProGlnAlaLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCCATCCACCGCAAAACAGCGCGTATGTTGACACTTACCCACTTCAGTACTG 960
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCCGACAGACTAATAATCTGGCAAAATCTCGGCGGCGCACTGAGACTCAACTG 1020
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheLysArgTrp 360
Db 1021 ACGCTTCCCGGTACACCGGATTAACAGCGCGCGAGGTGTGAACGTGTGTTGAACGCTGG 1080
Qy 361 ArgArgLeuSerAspAsnSerClnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

```

```

Db 1081 CCGCGCTAAGAGATTAACAGCCAGTGGATTCAGTTGCTGCTTCACAGACTTACAG 1140
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATCGCGTGAATAAACCGCGCTGATTAATTAATACCGCGCCGAGAGGTGAACGTGACC 1200
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGAGATGTGAAGAGCGAAATGCGCAGGCGATGTGTCTGTGGCAGTTTACGCA 1260
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCAGCATACCGCGCTGATTTGATGATCTCATCACCATCATCAC 1320

RESULT 3
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYLASES AND USES THEREOF
; FILE REFERENCE: DIVERI370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1998-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1

Alignment Scores:
Pred. No.: 4,1e-264 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-866-379-1 (1-1323)
Qy 1 MetLysAlaIleuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCGGTAAACCCGGAATCTGCA 60
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCCTCTGAGATGAGCCGCGAGCTGAAAGTGTGATGATGTCAAGTCGTCATGCT 120

```

OY	41	ValaIArgAlaIProThrLysAlaIThrGlnLeuMetGlnAspValThrProAspAlaITrpPro	60
Db	121	GTGGGTGCTCCAAACCCAAAGCCACGCACATGTATGCGAGATGTGCACCCCGAGACGACATGGCCA	180
OY	61	ThirTrpProValLysLeuGlnYTrrLeuThrProArgLysGlyGlnLeuIleAlaIatrrLeu	80
Db	181	ACCTGCGCGGTAAACTGGTGTGGCTGACACCGCGNGGTGGTGAAGCTAAATGCCCTATCTC	240
OY	81	GlyHisIstYrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
Db	241	GGACATTTACCAACCGCCACGCTGTGTATCCGACGAGTATGCTGGCGAAAAAGGGCTGCCG	300
OY	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluATrGThrArgLysThrGlyGln	120
Db	301	CAGCTGTGGTCCAGGTGCGCATTTATGTCTATGTCCGACGAGCGATACCCGTTAAACAGGCCAA	360
OY	121	AlaIleAlaIleAlaGlyLeuAlaIleProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	361	GCCTTTCGGCGCGGGCTGGCACCTGCATCTGTCATTAACCGTATACATACCAGGACAGATACG	420
OY	141	SerSerProAspProLeuPheAspProLeuIstYrThrGlyValCysGlnLeuAspAsnAla	160
Db	421	TCGAGTCCCGATCCGTTATTTAACTCTTAAAACTGGCGTTGGCAACTGATACACG	480
OY	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180
Db	481	AACGTAAGTACGCGGATCTCAGCAGGCGAGAGGATCAATTCGTGACTTTACGGCGAT	540
OY	181	ArgGlnThrAlaIlePheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys	200
Db	541	CGCGAAACGGCGTTTCGGCAACTGGAACTGGAGGGGTGCTTAATTTTCGCAATCAAACTTCGCG	600
OY	201	LeuIstYrArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu	220
Db	601	CTTAAGCTGAGAAACAGAGAGAAACGTTCATTAAACGACGAGCATTTACATCGGAATC	660
OY	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	661	AAGGTGACGGCGCAATGTCTCATTTAACCGGTGGCGTAAACCTTCGATCAATGCTGACG	720
OY	241	GluIlePheLeuLeuGlnGlnAlaGlnIleGlnIleMetProGlnIstYrTrpGlyArgIleThr	260
Db	721	GAGATATTTCCTCTGCAACAAACAGGAGATGCCGAGCGGGGTGGGGAAGGATCAACC	780
OY	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheIstYrLeuGln	280
Db	781	GATTCACACACAGTGGAAACACCTCTGTAAGTTTGCATTAACCGGCATTTTATTTGGCTACAA	840
OY	281	ArgThrProGlnValAlaIleArgSerArgAlaThrProLeuLeuAspIleIleMetAlaIa	300
Db	841	CGCACGCCAGAGGTGGCCCGCAGCGCGGCCACCCCGTTATTTGAGATTATCATGTGCACGC	900
OY	301	LeuThrProHisProProGlnLysGlnAlaIstYrGlyValIThrLeuProThrSerXAlaLeu	320
Db	901	TTTGACGCCCCATCCACCGCAAAACAGCGATGTGGTACATTACCCATCTTCAGTACTG	960
OY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp	340
Db	961	TTTTATTTGCGGACAGATATTAATCTGGCAATCTCGCGCGCCGCACTGTGACCTCAACATCG	1020
OY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp	360
Db	1021	ACGCTTCCCGGTACGCCGATTAACAGCGCCGACAGTGGTGAATAGTGCTTTGAAACGCTGG	1080
OY	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	1081	CGTGGCGTAAACGATTAACAGCCAGATGTGATGCAGTTTCGTGCTTCACAGACTTACAG	1140
OY	381	GlnMetArgAspLysIstYrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr	400
Db	1141	CAGATGGCTGTAAACCGCGCTGTCAATTAATACGCGCCGAGAGAGTAAACTGACC	1200

```

QY      401 LeuaIagIyCysGluGluaRgAsnaAlacInglwEtyCysSerLeuAIAglyPherthrln 420
Db      1201 CTGGAGCATGTGAAGACCGAAATGCCAGGCGCACTGTTCGTGGCAGGTTTAAACGCA 1260

QY      421 lIevaIasngIuaIaargIIeProalacYssSerleuAqgSerHISHisHSIsHS 440
Db      1261 ATCGTAATGAAGAAGCACGCATACC GGCGTGCAATTGAGATCTCATCCATCACCATTAC 1320

RESULT 4
US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
FILE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866, 379
PRIOR FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US 09/580, 515
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: US 09/318, 528
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: US 09/291, 931
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: US 09/259, 214
PRIOR FILING DATE: 1999-03-01
PRIOR APPLICATION NUMBER: US 08/910, 798
PRIOR FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patentin version 3.1
SEQ ID NO 7
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Alignment Scores:
Pred. No.:          6.62e-256           Length:         1901
Score:              2235.00             Matches:         430
Percent Similarity: 99.54%               Conservative:    0
Best Local Similarity: 99.54%            Mismatches:     2
Query Match:        97.09%                Indels:         0
DB:                 Gaps:                 0

US-09-777-566A-2 (1-440) x US-09-866-379-7 (1-1901)

QY      1 MetlysalaIaleIeuIleProPhelEuSerleuEuIleProLeuthrProGIInsEraIa 20
Db      188 AFGAAAGGAGCTTAATCCATTTTATCTCTTGCGATTGCGTTAAACCCGCAATCGCA 247

QY      21 PheaIacInserGIuProGIuLeuLysleuGIusEraValIalleIaSerARgHSIsGLy 40
Db      248 TTCGCTCGAGATGAGCCGCGAGACTGAGAAAGTGTGGATTTGTCAGTTCACGTCATCGT 307

QY      41 ValaIrgaIaAProThrlrlysaIaThrGlnueketGInaspValIThrProaPaIatTrpPro 60
Db      308 GTGGCGTCTCCAACCAAGCCAGCCCAACTGATGAGATGTGCACCCAGAGCGATGGCCA 367

QY      61 ThrTrpproVolvlsleuGIYTrrpleuThrProArGIyGIyGIuLeuIIaIATyrIleu 80
Db      368 ACCGAGCCCGGTAANAACtgggttgccgtgcACACCGcGNGStgGTgAGCTAATCGCCtATCTC 427

QY      81 GlyHIStryGINarGINarGInarGleuValaIaaspGIyeuLeuAaIatsLySGIcyasPro 100

```

```

|||||
428 GGACATTACCAACGCGCTCTGTACCCGCGATTGTCGGCAAAAAGGCGCCCG 487
|||
Qy 101 GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgIsthrGlyGlu 120
|||
Db 488 CAGTCTGCTCAGGTGCGATTATTCGTATGTCGCGAGCGCGTAAACCGAGCGCA 547
|||
Qy 121 AlaPheAlaIleAlaGluAlaProAspCysAlaIleThrValIstThrGlnAlaAspThr 140
|||
Db 548 GCCTTCGCGCGCGGCTGGCACCTGCTGCAATACCGTACATACCCAGCAATACG 607
|||
Qy 141 SerSerProAspProLeuPheAsnProLeuIsthrGlyValCysGlnLeuAspAsnAla 160
|||
Db 608 TCCAGTCCCGGATCCGTTATTTAACTCTAAACACTGCGTTGGCACTGATTAACGCG 667
|||
Qy 161 AsnValThrAspAlaIleLeuSerThrArgAlaGlySerIleAlaAspPheThrGlyHis 180
|||
Db 668 AACGTGACTGACGGCGATCTCAGCAGGCGAGAGGGTCAATTGCTTAACCGGCGAT 727
|||
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuGly 200
|||
Db 728 CGGCAACGGCGGTTTCGGCAACTGGACGGGCTTAATTTTCCCAATCAAACTTGTGC 787
|||
Qy 201 LeuIsthrArgGluIsthrGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
|||
Db 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATACGACGCAATTCATCGAGACGC 847
|||
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
|||
Db 848 AAGGTGACGGCGCAATGTCTCAATTAACCGGTGGGTCAATCGCATCAATGCTGACG 907
|||
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlyIsthrGlyValIsthr 260
|||
Db 908 GAGATATTTCTCTGCAACCAAGCACAGGGAATCCCGGAGCCGCGGGTGGGAGAGTCCAC 967
|||
Qy 261 AspSerHisGlnIsthrPasnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
|||
Db 968 GATTTCACACAGTGGAAACACCTTGCTAAGTTGCAATACGCGCATTTTATTTGCTACAA 1027
|||
Qy 281 ArgThrProGluValAlaIsthrSerThrArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
|||
Db 1028 CGCAGCGCCAGAGGTTGGCCGAGCGCCGACCCCGTATTAGATTGATCAAGACACGCG 1087
|||
Qy 301 LeuThrProHisProProGlnIsthrGlnAlaTyrGlyValIsthrLeuProThrSerValLeu 320
|||
Db 1088 TTGACGCGCCCATCCACCGCAAAAACAGCGGATGTGTGACATTAACCACTTCACTGCTG 1147
|||
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
|||
Db 1148 TTATATCGCGGACAGATACTAATCTGCAATATCTGGCGGCGGCACTGGAGCTCAACTGG 1207
|||
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValIsthrGluArgTrp 360
|||
Db 1208 ACGCTTCCCGGTTCAGCCGCGATACACGCGCGAGGTGTGAACCTGCTTTGAAACGCTGG 1267
|||
Qy 361 ArgArgLeuSerAspAsnSerGlnIsthrIleGlnValSerLeuValPheGlnThrLeuGln 380
|||
Db 1268 CGTGGCTTAACGCATTAACAGCAGTGTGATTCAGGTTTGGCTGCTTCCAGACTTTACAG 1327
|||
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValIsthrLeuThr 400
|||
Db 1328 CAGATGCGTGAATAAAACGCCCTGTCAATTAATACGCCGCGGAGAGTGAACATGACG 1387
|||
Qy 401 LeuAlaGlyCysGluGluIsthrAsnAlaIsthrGlyMetCysSerLeuAlaGlyPheThrGln 420
|||
Db 1388 CTGGCAGGATGTGAAGACGAAATGCGAGGCGCATGTTGCTTGGCAGGTTTATACGCA 1447
|||
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
|||
Db 1448 ATCGTGAATGAAGCAGCATACCGCGCTGCACTTTG 1483

```

RESULT 5
05-09-866-379-9

```

; Sequence 9, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYLASES AND USES THEREOF
; FILE REFERENCE: DIVERI370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Alignment Scores:
Pred. No.: 7,8e-255 Length: 1901
Score: 2226.00 Matches: 429
Percent Similarity: 99.318 Conservative: 0
Best Local Similarity: 99.318 Mismatches: 3
Query Match: 96.70% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-866-379-9 (1-1901)
Qy 1 MetLysAlaIleLeuIlePheProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
|||
Db 188 ATGAAGAGCATCTTAATCCCATTTTATCTTCTGATTCGCTTAACCCCGCAATCTGCA 247
|||
Qy 21 PheAlaGlnSerGluProGluLeuIsthrGlnIsthrValIleValSerThrGly 40
|||
Db 248 TTCGCTCAGAGTACCCGCGAGCTGAAGCTGGAAGTGTGATTGTGATGCTGATGTG 307
|||
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValIsthrProAspAlaTrpPro 60
|||
Db 308 GTGGCTGCTCCCAACCAAGCCAGCACGCACTGATGACAGATGTACCCAGACGCAATGCGCA 367
|||
Qy 61 ThrTrpProValLysLeuGlyTyrPleuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
|||
Db 368 ACCGTGCGCGTAAACTGGGTTGGCTGACACCGCGNGGTGTGACTTAATCCCTATCTC 427
|||
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIsthrGlyCysPro 100
|||
Db 428 GGACATTAACCAAGCCAGCGCTGTGTGACCGAGATTCCTGCGCAAAAAGGCGTGGCCG 487
|||
Qy 101 GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgIsthrGlyGlu 120
|||
Db 488 CAGTCTGCTCAGGTGCGCATTTTGTGATGTCAGACGACGCGTAAACCGAGCGCA 547
|||
Qy 121 AlaPheAlaIleAlaGluAlaProAspCysAlaIleThrValIstThrGlnAlaAspThr 140
|||
Db 548 GCCTTCGCGCGCGGCTGGCACCTGCTGCAATACCGGTACATACCCAGCAGATACG 607

```



```

|||||
Db 962 GATTACACACAGTGAACACCTTGCTAGTTGCAATACGGCCAAATTTATTACTACACA 1021
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
Db 1022 CGCACGCCAGAGGTTCGCCCGCAGTCGCCACCCCGTATTGATTTGATGCACGGC 1081
Qy 301 LeuThrProHisProGluProGluInuGluAlaTyrglyValThrLeuProThrSerValLeu 320
Db 1082 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTACATTAACCCACTTCAGTGCCTG 1141
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaIleuGluLeuAsnTrp 340
Db 1142 TTTATTGCCGCGACAGATACATATCTGCAAAATTCGCGGGCGGCACATGAGCTCAACTG 1201
Qy 341 ThrLeuProGlyGluProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1202 ACGCTTCCAGTCACCCCGATMACACGCCCGCAGGTGTGAACCTGGTTTGAACGCTGG 1261
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGluValSerLeuValPheGlnThrLeuGln 380
Db 1262 GCTCGGCTTAAGCGATTAACAGCCAGTGCATGAGTTGCTGCTTCACAGACTTACAG 1321
Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValuValLysLeuThr 400
Db 1322 CAGATGCGTGATAAACGCCGCTATCATTAATACGCCGCCGAGAGGTGAACACTGACC 1381
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1382 CTGGCAGAGATGTGAAGACGGAATCGCAGGCGCATGTGTCTGGCCGCTTTTACGCCAA 1441
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1442 ATCGTAATGAAGCCGCAATACCGCGCTGCAGTTTG 1477

RESULT 7
US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KREYFZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARRON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

```

```

Alignment Scores:
Pred. No.: 3 07e-254 Length: 1901
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 96.48% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-866-379-6 (1-1901)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGAGCATCTTAATCCATTTTATCTCTTCATTCCTGTAACCCGCAATGTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Db 248 TTCGCTACAGATGAGCCGAGGAGCTGAAGAGTGTGTGATGTCATGCTCATGTGT 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCTGCTCCAAACCAAGGCCACGCACTGATGACAGATGTCAACCCAGACGATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrgly 80
Db 368 ACTGCGCGGTAAACCTGGGTTGGCTGACACCGCGNGTGATGAGCTTAATCGCCTATCTC 427
Qy 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 428 GGCAATTAACCTGGGCCAGCGGTGTGTAGCGGAGATTCGTGCGAAAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 488 CAGTCTGTCAGAGTCCCGATATCTGCTGATGTGACGACGCTACCCGTTAAACAGCGGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTGCGCGCGCGGTGGCACCTGACTGTGCAATTAACGTTACATACCCAGGACGATAGC 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAACTCTTAACCAACGCGGTTGCCCACTGGGTAACGGG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTACTGACGCCGATCTCTACAGGCGAGAGGGTCAATTCCTACTTTACCGGGCAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCMAACGGCGTTTCGCGCACTGGAACGGGTCTTAATTTCCGCAATCAAACTGTGCG 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAGAAACAGAGCAAGACCTGTCAATTAAACGACCAATTAACATCGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTAGCGCCGACAAATGTCTCATTAACCGGCGGTAAAGCTCCGATCAATAGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTCTCTCTGCAACAGACAGCAAGGAATCCGAGACCGGGGTGGGGAAGATCACCC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnIleThrTyrglyLeuGln 280
Db 968 GATTACACACAGGAGCAACCTTGTTAAAGTTGTGCATTAACGCGCAATTTATTGTCATCAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
Db 1028 CGCACGCCAGAGGTTCGCCCGCAGTCGCCACCCCGTATTGATTTGATCAAGACAGCGG 1087
Qy 301 LeuThrProHisProGluProGluInuGluAlaTyrglyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTACATTAACCCACTTCAGTGCCTG 1147

```

QY	321	PhelllealaglyhsasptthAsmneualaasnleuglyglvalaleuGluLeuasntrp	340
Db	1148	TTTATTCGCCGACACGATACTTAACCTGGCAATATCTCGCGCGCACTGGACCTCAACATGG	1207
QY	341	ThreuprpglygnproaspasnthPrpproglyglgluleuValPhegluAargtrp	360
Db	1208	ACGCTTCCGGTCAACGCGGATPAACACGCCGAGGTGTGAACATCGTGTGTAACCGCTGG	1267
QY	361	ArgatgleuSerAspAsnSerGlntripleGlnValSerleuValPheGlntrhleuGln	380
De	1268	CGTGGCGCTAACCGATNAACAGCCAGTTCGATTCAGTTTGCTGGTCTTCCACACTTTTACAG	1327
QY	381	GlnMetArgAspLysThrProLeuSerLeuasnThrProProglyglValLysLeuThr	400
Db	1328	CAGATGCGTGATTAACCGCCGCTGCATTAATAACGCCGCCGCGAGAGGTAAACTGACC	1387
QY	401	LeuAlaGlyCysgluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Db	1388	CTGCACAGGATGTGAAGACGCAAAACGCCAGAGCGCATGTTCGTTGGCAGGTTTAAACGCA	1447
QY	421	IleValasnGluAlaArgIleProAlaCysSerLeu	432
Db	1448	ATCGTGATGAAGCACGCGATCCGCGCGTGCAGTTGG	1483

```

Db      1268  CGTCGGCTAACGATACACGACGAGTTCAGTTGCTTCAGACTTACAG 1327
Oy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db      1338  CAGATCGCTGATTAACCGCCCTGTCATTAAATACGCGCCGCGAGAGTGAACGTGACC 1387
Oy      401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388  CTGCGAGATGTGAAGACGAAATCGCAGGCGCATGTGTTCTGTCGACAGTTTACGCCAA 1447
Oy      421  IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      1448  ATCGGATGAAGCAGCATACCGCGCGTCAAGTTTG 1483

```

RESULT 9

```

US-10-021-723A-11
; Sequence 11, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Yersinia pestis
US-10-021-723A-11

```

```

Alignment Scores:
Pred. No.: 1,986-104      Length: 1326
Score: 960.50             Matches: 202
Percent Similarity: 61.24% Conservative: 65
Best Local Similarity: 46.33% Mismatches: 154
Query Match: 41.72%      Indels: 15
DB: 9                      Gaps: 7

```

US-09-777-566A-2 (1-440) x US-10-021-723A-11 (1-1326)

```

Oy      5  LeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSer 24
Db      37  TTAGCTGTAATGCTCTTCGGGCTGGCTGCTATCACCGCGCCGTGTGCC-----GCC 87
Oy      25  GluProGlu---LeuLysLeuGluSerValIleValSerArgHisGlyValArgala 43
Db      88  GAACCATGCGGGGTACACCTTAGAACGTGATTTAGTATTTAGAGCGCCATGGTGTGGTACC 147
Oy      44  ProThrLysAlaThrGlnLeuMetLysPheValThrProAspAlaThrProThrTrpPro 63
Db      148  CCGACTAAGCAGACACGACCTGATGATGATGATACACCTGATMACTGGCTCAGTGGCCG 207
Oy      64  ValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeuGlnLysTyr 83
Db      208  GTTAAAGCGGGCTATTACTCTCTGTCGCGCGCAACTGGGTACCCCTGATGGCGGGTTTC 267
Oy      84  GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProGlnSerGly 103
Db      268  TATGCGCATATTTCCGAGATTGGTCTTTTGGCC---GCGGGGTGGCCCGCAGAGGGC 324
Oy      104  GlnValAlaIleIleAlaAspValAlaSpGluArgThrArgLysThrGlyGluAlaPheAla 123
Db      325  GGTGTATATGACACAGACAGATATGACACAGCGACTCGTTTAAACGGGTACAGCTTTTCTG 384

```

```

Oy      124  AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerPro 143
Db      385  GATGGTGTGCGCCCGGGTGGCGGCTGACGTGCTGCCAATTCACAGCCGATCTTAAGAAACC 444
Oy      144  AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr 163
Db      445  GATCCTCTGTTTCATCCCGCTGTAACCCGCGTGTGAATCATGTGACAAACCCCAACCCGAT 504
Oy      164  AspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGlnThr 183
Db      505  AAGCAATTTGAGGAACCGCTGGCGCGCCGTAGACACCGTAAACCGCATGCGTATGCCAAA 564
Oy      184  AlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCysLeuArg 203
Db      565  CCGTTTGGCGCAATATGGCGCATGCTCTGAACCTTCGTCGACAGTCCGACTGCAAGTCACAG 624
Oy      204  GlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
Db      625  CAGCAGACAGAGGGAACAACTTGACTTCGCACACTTTCGCGCCACAGAAAGTTAATGTAAAC 684
Oy      224  AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlu 241
Db      685  AAGGAAGGCACGAAAGTTACCTGTCAGGCGCCCGCTGCTAGACACGCTTGAGCGGCA 744
Oy      242  IlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPolYArgIleThrAsp 261
Db      745  ATCTTCTTCTGTCAGAACCGCAGCGAGTCCCGCAAGTACGCTGCGACGCTTTGAAAGGC 804
Oy      262  SerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArg 281
Db      805  GCTGAGAACTGGGTGTCCTCTGAGCCCTGCACAAATGCACAGTTCACACTGATGGCTA 864
Oy      282  ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeu 301
Db      865  ACCGCATACATTGCGACCGCAACAAAGCAGCGCTTTTACACCAATATCATACCGCATCG 924
Oy      302  ThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer----- 318
Db      925  ACC-----CTGCAACTGAGACCGCCCGGCGGCAAAACCTGCCGATCTCGGCTCAGAAC 975
Oy      319  ---ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGln 337
Db      976  CGTGTATTATTCCTGGTGGCCGACGACCAAAATATGCTAACATCGCCGCTATGCTGGGCG 1035
Oy      338  LeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPhe 357
Db      1036  GCAGATGGCAGTTACCGGAACACCGGATTAACACCCCGCGGCGGCTGTGCTTT 1095
Oy      358  GluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGln 377
Db      1096  GAGCTGTGCGCAAAATCCGCAATATCATCAACGTTATGTGCGGCTTAAGATGTTATCAG 1155
Oy      378  ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeu---AsnThrProProGlyGln 396
Db      1156  ACCATGATGATCAATGGGTGAACCGCGAGAGCTGATTTAAAGAACAAATCCCGCGGCATC 1215
Oy      397  ValLysLeuThrLeuAlaGlyGluGluArgAsnAlaGlnGlyMetCysSerLeuAla 416
Db      1216  ATCAGTGTGCTGTGGCGCGCTGCGGATTAATGCTGACGATTAATGTGCAACATTGAT 1275
Oy      417  GlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      1276  ACTTTTCAAAAAAAGTTAGCGAAAGTCATTGAACCTGCTGTCAATATT 1323

```

RESULT 10

```

US-10-021-723A-3
; Sequence 3, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan

```

```

; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1325
; TYPE: DNA
; ORGANISM: Versinia pestis
; US-10-021-723A-3

```

Alignment Scores:

```

Pred. No.:      8,24e-102      Length:      1325
Score:          938.50         Matches:      202
Percent Similarity: 61.24%      Conservative: 65
Best Local Similarity: 46.33%      Mismatches:  154
Query Match:    40.77%         Indels:       16
DB:             Gaps:         7

```

US-09-777-566a-2 (1-440) x US-10-021-723A-3 (1-1325)

```

QY      5 LeuIIePProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSer 24
DB      37 CTGGTATTATGATCGATGCGATGCGCTGCTATACGCGCGGTACCC-----GCC 87
QY      25 GlnPProGlu---LeuysLeuGluSerValIleValSerAlaThrGlyValAlaGln 43
DB      88 GACCCATCGCGGCTATCTTACGCTGAGCTGTTGAGTGCCTGCGCATGGTGTCTGCG 147
QY      44 ProThrLyAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThrTyro 63
DB      148 CCAGCAACAAACA-ACGAGGTTATGATGATGATGATGATGATGATGATGATGATGATG 206
QY      64 ValysLeuGlyTyrrLeuThrProAlaGlyGlyGlyLeuIleAlaThrLeuGlyHisTy 83
DB      207 GTAAAGCGGGGATTTAAAGCCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCT 266
QY      84 GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProGlnSerGly 103
DB      267 TATGTCGATTAATCTTTCGACGCTTGGTTGTTAGCG---CGGGAGATGCGCGAGAGG 323
QY      104 GlnValAlaIleIleLeuAlaAspValaAspGlyArgGlyThrGlyGlyAlaPheAla 123
DB      324 GGGGTATATGACACAGCAGATATGATGATGATGATGATGATGATGATGATGATGATG 383
QY      124 AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSer 143
DB      384 GATGGTGTGGCTCCGGGGTGTGGTGTGACCGTGCATCAATCAGCCCATTTGAAAAAG 443
QY      144 AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr 163
DB      444 GATCCCTGTTTCATCCGTAAGACTGGCGTGTGTAAGTTAACAACGACAAACAGAT 503
QY      164 AspAlaIleLeuSerAlaGlySerIleAlaAspPheThrGlyHisArgGlnThr 183
DB      504 AAAGCATTTGAAGAACGATTTGGCGGCGCTGATGATGATGATGATGATGATGATG 563
QY      184 AlaPheArgGlyLeuGlnValAlaLeuAsnPheProGlnSerAlaLeuLysValArg 203
DB      564 CCTTTTGGCCAGTGGGAGCGTGAATTTGGCGTGTCTTATTCGAATCTTTG 623
QY      204 GlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGlyLeuLysValSer 223
DB      624 CAACAGCAAGAAAAAGTGTGATTTTCCCGCATTTGGCGCCCAATGATTAAGTTAAT 683
QY      224 AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlu 241
DB      684 AAAGAAAGGACAAAAATGACCTCAGTGGCGCCGCTGATATCATCATGACATTTGGGTGA 743

```

```

QY      242 IlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrrPheArgIleThrAsp 261
DB      744 ATCTTATTATTACAAAAGCAGCAAGCATGCCAGATGCTCGGCAACGGCTAAAGGG 803
QY      262 SerHisGlnThrAsnThrLeuSerLeuHisAsnAlaGlnPheThrLeuGlnArg 281
DB      804 GCGGAGAAATGGGATCATTTATGTCATATACATACGCGCAATTTAATTAAGCAAAA 863
QY      282 ThrProGlnValAlaArgSerAlaThrProLeuLeuAspPheLeuAlaLeu 301
DB      864 ACACCGTATATGCCCGCTGCTATTAAGAGGCCCATTTTACAGCATATGATGAGCTTTA 923
QY      302 ThrProHisProGlnLysGlnAlaTyrrGlyValThrLeuProThrSer----- 318
DB      924 ACC-----CTTCAACGATGCTCAGAGGGCAAAAGCTATTCAGCCCAAAAC 974
QY      319 ---ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlu 337
DB      975 CGGGTCTTGTCTCCGTCGGTGGCAATGATACCAATATTTGCAATTTGCGGGTATGCGGA 1034
QY      338 LeuAsnThrProLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPhe 357
DB      1035 GCGGACTGGCAGGTACCCGACGACCACTGATATACACACAGATGGGGGATTTGTTTTT 1094
QY      358 GlnArgTrpArgArgLeuSerAspAsnSerGlyTrpIleGlnValSerLeuValPheGln 377
DB      1095 GAACATATGGCAAAATCCGATATACACACAGCCGCTAGCTGCTGAAGATGTTCTACCA 1154
QY      378 ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeu---AsnThrProProGlyGlu 396
DB      1155 ACAGATGATCAGTGGCTATATGCGGAAATTTGATCTGAAATAATACCAAGCGGATATT 1214
QY      397 ValysLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnLysMetCysSerLeuAla 416
DB      1215 ATTCCGTCGACATGCTGCTGTGTAATAATACGATGACGATGATGCTTGGCAGCTTGAT 1274
QY      417 GlyPheThrGlnIleValAsnGlnAlaAlaArgIleProAlaCysSerLeu 432
DB      1275 ACATTTCCAAAAGAAAGTGGCTAAGTAATTTGAACCTGCTGCCACATC 1322

```

RESULT 11

```

US-10-021-723A-1
; Sequence 1, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1325
; TYPE: DNA
; ORGANISM: Versinia pestis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 470, 487, 491, 492, 522
; OTHER INFORMATION: n = A,T,C or G
US-10-021-723A-1

```

Alignment Scores:

```

Pred. No.:      1.98e-99      Length:      1325
Score:          918.50         Matches:      199
Percent Similarity: 60.32%      Conservative: 64
Best Local Similarity: 45.64%      Mismatches:  158

```

OY	5	LeuileProPheleuSerleuLeuileProleuThrProglInserAlaPheAlaGlnSer	24
Db	37	CTGGATTGATGCTTAAAGCGGATTTGGCTGCTATTACGCGCCGGATACC-----GCC	87
OY	25	GluProGlu---LeuLysleuGlnSerValAlaIleValSerArgHisGlyAlaArgAla	43
Db	88	GAGCATGGGGCTATACCTTATAGACGTGTGGTATTATTGAGTGGCCATGGTGTTCGCTCG	147
OY	44	ProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThrTTPPro	63
Db	148	CCGACCAAAACAAACCCAGCTTATGAAATGATGTTACCCAGATGAATGAAGTGGCCAAATGGCCG	207
OY	64	ValLysleuGlyThrLeuThrProArgGlyGlyGlnLeuIleLeuIleAlaThrLysHisTyr	83
Db	208	GTAAGAAGGGGGTATTTAACGCCACGTGGTGGGAACTTGGTCACATGATGGGGGGTTT	267
OY	84	GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSergly	103
Db	268	TATGGTGAATTACTTTCGACGCTGGTGTGTGTAAGC---CGGGAGTGTCCGGAGAGGGG	324
OY	104	GlnValAlaIleIleAlaAspValAspGluArgGlyThrArgLysThrGlyGlnAlaPheAla	122
Db	325	GGGGATATATGCACAGGACAGATATCGATCAACGTACCCGCTTAACCGGACAGGATTTCTT	384
OY	124	AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSergSerPro	143
Db	385	GATGGTGTGGGCTCGGGGGTGTGGTTGACCCGTGCATATCAGGCCGATTTGAAAAAGACC	444
OY	144	AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr	163
Db	445	GATCCCTCGTTTCCATCCGGTAGAGTGGCGTGTGAATTAATTAACGNNGCACAAACAGAT	504
OY	164	AspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlnHisArgGlnThr	183
Db	505	AAAGCGATTGAAGAACAMTTGGGGGGCCGTTAAGATACGGTGGCCAGCGCTACGCTAAA	564
OY	184	AlaPheArgGlnLeuGluArgValLeuAsnPheProGlnSerAsnLeuLysLeuLysArg	203
Db	565	CCTTTTGCCACAGATGGGGACGCTGATTTTGGCGCTTCCTCTTATTTGCAATCTTTG	624
OY	204	GluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSergLeuLysValSer	223
Db	625	CAACACGCAAGGAAAAACGTGATTTTGGCCACTTTCGGCCCAAT--CAAGCTTAACGTTAAT	683
OY	224	AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerLeuLeuThrGlu	241
Db	684	AAAGAGGAGCAAAAGTGACCTCAGTGGGCCACTG6GCTTATCATGACGACATTGGGTGAA	743
OY	242	IlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyThrPylArgIleThrAsp	261
Db	744	ATCTTCTTATTAACAAACGACACACCAACGACGAGAGGTTCGCTGGCAAGCGCTAAAAAGG	803
OY	262	SerHisGlnThrPAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGlnArg	281
Db	804	GCGGAGAAATGGGTATCCTTATTTGCATTTACATAACGCCCAATTTATTTAAAGGCAAAA	863
OY	282	ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIleLeu	301
Db	864	ACACCGTATATCGCCGCTATMAAGGAGCCATTTATTCACAGCAGATAGATACGGCTTTA	923
OY	302	ThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer-----	318
Db	924	ACC-----CTTCAACTCGATGCTCAGGGGCAAAACCTACCCATTTCAGCCCAAAAC	974
OY	319	---ValleuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuLysGlyAlaLeuGlu	337
Db	975	CGGGTCTTGTCTCTCGGTGGGATATACCAATATTTGGCCAAATATTCGGGGTATGCGGA	1030

```

QY      338  LeuAsnTrpThrLeuProGlyInProAspAsnTrpProProGlyGluLeuValPhe 357
      : ::::: ||||| ||||| ||||| ||||| ||||| |||||
Db      1035 GCCGACTGGCACCTGACCCAGCAACCTGATTAATACCTCACAGATGGGGATGGCTTTT 109
QY      358  GluAgtTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGln 377
      ||| ||::: ::::: ::::: ::::: ||| ::::: :::::
Db      1095 GAACTATGGCAAAATCCGATAAACCAACCAAGCTACGTTGCGGTGAAGATGTTCTACAA 115
QY      378  ThrLeuGlnGlnMetArgAspLysTrpProLeuSerLeu---AsnTrpProProGlyGlu 396
      ||::: ||::: ||::: ||::: ||| ||| ||| |||
Db      1155 ACGATGATCAGTGGCTGATGGCGCAAAATATGGATCTGAAATAATACCAAGGGGATAT 121
QY      397  ValLysLeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnIleCysSerLeuAla 416
      ::::: ||||| ||||| ||||| ||||| |||||
Db      1215 ATTCCTGGTGCAGTTCGTGGTGTGAAATTAACGGTGACATTAAGCTTTCGAGCTGAT 127
QY      417  GlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 432
      ||| :::: ||| :::: ||||| |||||
Db      1275 ACATTCCAAAGAAAGTGGCTAAGTAATTGAACCTGCTGCCACATC 1322

RESULT 12
US-10-021-723A-9
; Sequence 9, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016C1P)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 954-2 Phytase sequence
US-10-021-723A-9

Alignment Scores:
Pred. No.: 1,63e-80 Length: 1431
Score: 760.00 Matches: 166
Percent Similarity: 54.55% Conservative: 80
Best Local Similarity: 36.81% Mismatches: 177
Query Match: 33.01% Indels: 28
DB: 9 Gaps: 9

US-09-777-566A-2 (1-440) x US-10-021-723A-9 (1-1431)
QY      4  IleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSer---AlaPheAla 22
      ::||::: ||| ::||| ||| :::: |||
Db      64  GATTGCTTTGGCGATGCTCAACGCAACTGCGGGCTGCTCAACGCTGTGCGCGG 123
QY      23  GlnSerGluProGlnLeuLys-----Leu 30
      ||||| |||
Db      124  GGTACGACGACCGCCCGGAAAGCCGCGGCAACAGATGGCATGCCGCAAGACTGTGGCTC 183
QY      31  GluSerValValIleValSerArgHisGlyValArgAlaProThrLysAlaThrGlnLeu 50
      ::::: ||||| ||||| ||||| ||||| |||||
Db      184  GACGGGTGGTCATGCTCAGCGCGGACGCGCTGCGGCTCTCCGACGCGCTCCGAGCCGCG 243
QY      51  MetGlnAspArgValThrProAspAlaTrpProThrTrpProValLysLeuGlyTrpLeuThr 70
      ::::: ||||| ||||| ||||| |||||
Db      244  CTGGAGACCTCTACGCCGCGATCCGTGGGCGCCAGTGGCCCTGCGCACTGCCACCTGACC 303
QY      71  ProArgGlyGluGluLeuIleAlaTrpLeuGlnHisTyrGlnArgGlnArgLeuValAla 90

```

```

Db      304 GATCGTGGCGCGCGCTGCTCGCAGATGGGGGCTACTACGGTATGCTGCTGCTGCC 363
Oy      91 AaPGLyLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaasp 110
Db      364 GGGGGTGTCTCCCGCCGACGCGGGTCCCTCGGACCGGAAACCTTTACGATGGGACAGAC 423
Oy      111 ValaSPGLuArgThrArgLysThrGlyGlnAlaIleAlaLysLysLysLysLys 130
Db      424 GTTGCACAGCGGAGCGTGTGACGCGGACCGCTCTCCGTCCGTGACGCGGACGCTGCC 483
Oy      131 AlaIleThrValHisThrGlnAlaAspThrSerProAspProIleuLeuAsnProIleu 150
Db      484 GGTATCCACAGGATCATCGCGCGGCTCGACAGAGAACGATCCGATCTTCACAGCGGATG 543
Oy      151 LysThrGlyValLysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAla 170
Db      544 GAATGGGGTGCATGGCCAGTCGACCCCGTACAGCGGACGCGGACGATCGAACGCGATGCC 603
Oy      171 Gly---GlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgLysLysLys 189
Db      604 GCGGAGGGGCGCGTGGCGACACTGGGAGAGCGCTACGACAGCAAGCTTGACAGAAATGAGC 663
Oy      190 ArgValLeuAsnPheProGlnSerAsnLeuCysLysLysArgGlnLysLysLysLysLys 209
Db      664 GAGGTGCTGACTACGCCCATAGCCGCGCATTTGCGGAGGCGATGGCGCCAA----- 714
Oy      210 CysSerLeuThrGlnAlaLeuProSerGlyLeuLysValSerAlaLys-----AsnVal 227
Db      715 TCGGACTATGGCGGCCAA---CCCAATGCTGTGCAGATGACAGACAGATGGCTTCAATGCC 771
Oy      228 SerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLysGln 247
Db      772 GCGTTGAAGGGCCGACAGTGGCGAGTCTGACCGCTTCGACGCTCTCCGTCTCGCAACAT 831
Oy      248 AlaGlnLysMetProGlnPro-----GlyTrpGlyArgIleThrAspSerHisGlnTrp 265
Db      832 GGGGACAGGGCTCCACAGAGACAGGTGATGGGCGGATATCCACAGATCGCAGACAGCTGG 891
Oy      266 AsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGlnArgThrProGlnVal 285
Db      892 ACGCTGCTGATGACAGCGCATGAAAGCGGAGTTGATGATGATGCGGAAAGCGCTTACATG 951
Oy      286 AlaArgSerArgAlaThrProLeuLeuAspIleMetAlaIleAlaLeuThrProHis--- 304
Db      952 GCTACTCGAAGGGGACCGCATGCTCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1011
Oy      305 -----ProProGlnLysGlnAlaValTrpGlyValThrLeuPro-----ThrSerVal 319
Db      1012 GCGGCGCCAGCTCCGAGCTTCCGCTCAAGGCCCGGACGCTGCCCAAGGCCAAGCGTGTTC 1071
Oy      320 LeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyLysAlaLeuGlnLeuAsn 339
Db      1072 TATGTGCTGACCGCGCATGACACGAAATCTTGCAACTTGGCGCGGTGGTGGACTGTGAC 1131
Oy      340 TrpThrLeuProGlyGlnProAspAsnThrProProGlyLysLysLysLysLysLysLys 359
Db      1132 TGGACCTCGCCGACACCGGACGACGCGCGCGGCGGTGCATGTTGTTCTCTCTG 1191
Oy      360 TrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeu 379
Db      1192 TGGCGGAGCGCGGACGACGACGCTTCCGCTGCGGTGGATGGTCTATACATGATGATG 1251
Oy      380 GlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyLysLysLysLys 399
Db      1252 GATCAGCTTCGGACACTACGCGCGCTCCCTGGGCGAGCGCGCCCATGGCTGATCTTG 1311
Oy      400 ThrLeuAlaGlyCysGlnGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThr 419
Db      1312 CCGTTGGCCGCGTGTGGACGCGCGGACGACGATGATGATGATGATGATGATGATGATG 1371
Oy      420 GlnIleValAsnGlnAlaArgIleProAlaCys 430

```

```

Db      1372 CCGCGTGTGCGCGCGGCGCATTTGCCCTCTCTGCC 1404
RESULT 13
US-10-021-723A-7
; Sequence 7, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Rhizobium
; US-10-021-723A-7

Alignment Scores:
Pred. No.: 7, 27e-47
Score: 476.50
Percent Similarity: 45.548
Best Local Similarity: 31.358
Query Match: 20.708
Gaps: 15

US-09-777-566A-2 (1-440) x US-10-021-723A-7 (1-1230)
Oy      10 SerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerLeuPro----- 26
Db      10 AGCTGCTTTTGTGCTGCTCCGCGCTGCGGCGGATGCGCGCGCTCCGCGGAGCGGAC 69
Oy      27 -----GluLeuLysLysGlnLysValValValIleValSerArgHis 39
Db      70 GCATGCGCTCCGCGCGCGCGGCGGCTGCTGAAAGCTCGAAAGAGTGTAAATGCTGATGCCGAC 129
Oy      40 GlyValAlaArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAlaThrProAspAlaTrp 59
Db      130 GCGGTGCGCGCGCAACGAGGCGGCGGTGCGCGCGCGGTATTCGCGCGGAAACATGG 189
Oy      60 ProThrTrpProValLysLeuGlyTrpLeuThrProArgGlyLysLysLysLysLysLys 79
Db      190 CCCGACTGGCGCGTTCATTTGCGCTGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCTT 249
Oy      80 LeuGlyHisLysLysGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysLys 99
Db      250 CTGCGGGAAGGACCGCGCTCTATTTCCGCGGCTCGCGGCTGTT---CCCGAGGATGC 306
Oy      100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspValArgThrArgLysThrGly 119
Db      307 CCGGCGCGGCGACGATGCTCTCAAGGAGGACGACGAGGCGGCGGCGGCGGCGGCGGCGG 366
Oy      120 GlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db      367 CAGAGCTGGCGCGGCGGTTCAATGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGATGC 417
Oy      140 ThrSerSerPro-----AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeu 157
Db      418 CCCGCGGCTCCGAGACGATGACGCGATCTTCAATGGCTCGACGCGGCGCGCGCGCTTC 477
Oy      158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAla-----GlyGlySerIleAla 175
Db      478 GACGGCAAGCGGCGGTGATGCGCGGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 537
Oy      176 AspPheThrGlyHisArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPhePro 195

```

```

Db      538 GAAACCGACGTCACGCGC---GGCGACTGACCTTCTCGCGCAAGATGCTGAAAT-----588
Oy      196 GlnSerAsnLeuCysLeuLysArgLysGlnAspGlnSerCysSerLeuThrGln----214
Db      589 -----TGGCGCTGCGCCCGCTGC
Oy      215 -----AlaLeuProSerGluLeuLysValSerAla---AspAsnValSerLeuThr230
Db      607 CCGCGATCGCGAGCGGAGCGGCTGTGTCGGCGGACCGCGACGATCGCCCGCAGATCGGA666
Oy      231 G1yAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGly250
Db      667 GCGCGGCTCGAGCTGCGATCGACCGCCAGACGACGCTGCTGGAATATCTGGAAGGC726
Oy      251 MetPro-----GluProGlyTyrProLysArgIleThrAspSerHisGlnTyrAsnThrLeu268
Db      727 AAGCGGATGGCGAGGTGGCTGGGGCGCGCGTGAAGCCCGCC---GAGATCGAGCAGATTG783
Oy      269 LeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluValAlaArgSer288
Db      784 CTGGCGCTTCATCCCGCTCAAAATTCGGCTATTCGAAATCGCCCGCGTATATCGCCGACAC843
Oy      289 ArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProProGlnLys308
Db      844 GCGCGGCGCGGATCGTGGCGAATCGTCACGCGCGCTC-----882
Oy      309 GlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsn328
Db      883 -----GGCGACCGAAGCCCGCGCGCGCTGACCTTGTCTCGCGGCGACACAGAAC933
Oy      329 LeuAlaAsnLeuGlyAlaLeuGlnLeuLeuAsnTyrProThrLeuProGlyLysPro---Asp347
Db      934 GTCGCGGACCTCGCGGCTTCTTCTGACCTGCGACCTGCGACGATCCGAGTTATATCCCGCGAC993
Oy      348 AsnThrProProGlyGlyLeuValPheGlnArgTyrArgArgLeuSerAspAsnSer367
Db      994 GAGTTCGCGCGCGGACGCGACCTGGGTTGAA-----CTGGTCAGCAATGGC1041
Oy      368 -----GlnTyrPheGlnValSerLeuValPheGlnThrLeuGlnGlnMetArgAsp384
Db      1042 AAGCGGACCGCTATGCTCGCGCTTCATTCGCGGCGAGACGATGACCGACGCTCGCGAAC1101
Oy      385 LysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGlyCys404
Db      1102 CTCGACGCGCTGAGGTGCGACGATACGCTGTTCGCCCGCTATCTCCCATTCGCGGGTGC1161
Oy      405 GluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIle421
Db      1162 GGCATATCGCTCGAGCGACGCGCTGCGCTGAGTGAATTCGCGCGGCTC1212

RESULT 14
US-10-021-723A-5
; Sequence 5, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Mathur, Jay
; APPLICANT: Short, Jay
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: 953-6 phylase sequence
US-10-021-723A-5

Alignment Scores:
Pred. No.: 2,27e-46 Length: 1266
Score: 472.50 Matches: 137
Percent Similarity: 44.55% Conservative: 59
Best Local Similarity: 31.14% Mismatches: 177
Query Match: 20,538 Indels: 67
DB: 9 Gaps: 15

US-09-777-566a-2 (1-440) x US-10-021-723A-5 (1-1266)
Oy      10 SerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro-----26
Db      46 AGCTGTCTTTTGTTCGCGCGCTGCTGGGGGATGCGCGCTCGCGCGCGCGCGAC105
Oy      27 -----GluLeuLysLeuGlnSerValValIleValSerArgHis39
Db      106 GCATCGCCCTCGCGCGCGCGGGGTGCGTAAAGCTCGAAAGGCTGCTGATCGTACGCCAC165
Oy      40 GlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyr59
Db      166 GCGGTTCGCGCGCGCAAGCGCGGTGTGCGCGCGGTATATCGCGCAAAACATCG225
Oy      60 ProThrThrProValLysLeuGlyTyrPheLeuThrProArgGlyGlyLeuLeuIleAlaTyr79
Db      226 CCGGACTGGCGGTGATTTGCGCTGCTGACGCGCGACGCGCGCGGGGGGTCAAGCTG285
Oy      80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys99
Db      286 CTCGCGGAAGGACCGCGCTCATCTTGGCGGGTGGCGGCTATTC---CGGACGGTTGT342
Oy      100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspLysArgThrArgLysThrGly119
Db      343 CCGCGCGCGGACGATCTGCTCAAGACCTACAAAGACGCGACGATCGCGACGCGG402
Oy      120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp139
Db      403 GAAACTGGCGCGCGGTTCATTCGCGGCTGCACG-----GGCGAT444
Oy      140 ThrSerSerPro-----AspProLeuPheAsnProLeuLysThrGlyVal154
Db      445 GTCGCGGATCCGCGCGGTGCGGACGATGACGGATCTTCATGCGGTCCACGCGGCGCC504
Oy      155 CysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAla-----GlyGly172
Db      505 GCGTGTTCGACGCGGACGCGGATTCGATGCGCGCTGCGCGGCGCGGCGG564
Oy      173 SerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValAla192
Db      565 CTGACCGCGGAAACCGCACCGCATCGC---GGCGAATGACCTTGTGCGGAAAGTGTG621
Oy      193 AsnPheProGlnSerAsnLeuLysLeuLysArgGluLysGlnAspLysSerCysSerLeu212
Db      622 AAT-----TGTGGCTG633
Oy      213 ThrGln-----AlaLeuProSerGluLeuLysValSerAla---AspAsnVal227
Db      634 CCGCGCTCGCGCTGATCGCGCGCGCGGCTGCGCGGCTGCGCGGCGCGCGCGCGCGCG693
Oy      228 SerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGln247
Db      694 GATCTGGAAGCGCGCTGCGATCGGATGACCGCGGACGACGACGCTGCTGCTGGAATAT753
Oy      248 AlaGlnGlyMetPro-----GluProGlyTyrGlyArgIleThrAspSerHisGlnTyr265
Db      754 CTGGAAGCAAGCAGATGCGGATGCGGTGCGGTGAAGCGCGCC---GAGATC810
Oy      266 AsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGluVal285
Db      811 GAGCAGTTGCTGCTTCATTCGCTCAAAATTCGCTATTCGAATCGCGCGCGCTATATC870

```



```

Db      1225  |||::|||      ::|||  |||||  ::|
          ACATGCAATGTAAGTCAGAAAGAACCTTTA---GTTAGAGCTTTGATTAAACGACAGA 1281
OY      398  LysLeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGly 417
          ::|  |||  |||||::|  ::|  |||  |||  |||
Db      1282  GTTGTCCATTACATGGTTGTGATGTT---GACAAAGTTAGGTAGATGTAAGTTAAACGAC 1338
OY      418  PheThrGlnIleValAsnGluAlaArg 426
          |||  ::|  ::|||  |||||
Db      1339  TTTGTCAAAGGCTTGTCTTGCGCCAGA 1365

```

Search completed: June 12, 2003, 13:21:12
 Job time : 275 secs

THIS PAGE BLANK (USPTO)